Towards a Framework Using Sonic Expression as a Gaming Mechanic for Immersive Playing Experiences in VR Games

Art & Computational Technology PhD

Goldsmiths, University of London

Alexander Brigden

Student No. 33313034

Acknowledgements

Firstly, I would like to express my sincere gratitude and appreciation for my patient and supportive supervisor Professor Mick Grierson. It has been a long journey, but we have finally made it to the finish line! There are far too many people (past and present) from the Goldsmiths Computing Department for me to list, so many of you have been incredibly supportive, especially in times of need. I look forward to celebrating with you all. I am thankful beyond words to my friend Dr Emily West, for your proofreading, stats chat, and general guidance over the years. Your assistance was invaluable. An honourable mention to two of my fellow RHB329 peers; Dr Christian Guckelsberger and Rob Homewood, both of whom were incredibly helpful and enthusiastic towards my research in the early years at Goldsmiths. I extend my heartfelt thanks to all my friends who have resisted the urge to ask me 'So, how is the PhD going?' And a special mention to those who put me up when I first moved to London. You were incredibly accommodating and helped make this all possible. 'Grove Street for life!' Last, but certainly not least, I am so truly appreciative of the support from all my family. Whether it was games of Yaniv, mountain escapes, or an indulgent late lunch (accompanied by two cheeky monkeys, both of whom did not exist when I started!), you have always been there to provide comfort, support, and a much-needed respite from the trials and tribulations of a PhD. I am so incredibly lucky; this would not have been possible without you all behind me.

Abstract

This practice-based thesis addresses how sonic expression can be used within game design to make immersive, engaging playing experiences. The use of sound in games immersion is often a passive response to the player's in-game interaction. Could this immersion be increased by offering the player a broader range of interaction methods for sound expression? Would this alter the overall playing experience of the game? By creating games with novel game challenges that incorporate inventive in-game sound interaction methods, this research studies the impact that sonic expression has on the playing experience. We particularly focus on player immersion, determining whether increasing the depth of sonic expression produces measurable change. Screen-based and Virtual Reality platforms are compared to determine the effect of sound interactions within each. How player agency on sound sequencing and game challenge navigation affects agency in the playing experience is also explored. Different sound-design approaches are tested, comparing the use of diegetic and non-diegetic sound and its impact on synchronised player performance and experience. By studying the relationship between sonic expression and its impact on the game challenge and player immersion, we can inform game design practice. This leads to novel game and sound design approaches, highlighting potential new methods for generating procedural content relating to sound interactions. These findings are codified as a conceptual framework, which provides guidelines for sonic expression in game design. The DIVE (Demand, Inclusivity, Versatility, Engagement) framework's four domains highlight the performative nature of sonic expression, how sound frames the in-game challenge, the openness of sound-challenge space, and the continuity of the use of sound when designing games. Following these guidelines will lead to better immersion and a positive impact on the overall playing experience. This research benefits the field of game design research while also providing much-needed guidance for the use of interactive sound game design practice.

Table of Contents

List of Tables

List of Figures

Introduction

Background

Bleep. Bloop. Bleep… From the first analogue circuit generating sound in Computer Space (Nutting Associates, 1971), sound has been a component in video games. Whether synthesised by electronic circuits or played from an audio file, sound has been used to enhance the playing experience (Grimshaw, 2014). As processing power has increased for gaming devices, game designers have been able to employ a variety of Digital signal processing (DSP) techniques and effects to situate the player within the game world. Despite the ever-increasing technological affordance (Thubron, 2020) of these gaming devices, interactivity in sound has largely been constrained within paradigms originating in film (Alves & Roque, 2011).

By looking at existing methods of sound interaction found within electronic music production, we can identify appropriate methods and techniques that have the potential for a novel approach, offering new and engaging experiences for the player. With the use of synthesis, sequencing, and DSP techniques such as filtering, this approach changes sound from a passive response to an active prominent component of the game. This creates a potential opportunity for sound to be used expressively by the player, a method to communicate (Fels et al., 2002) within the sound domain that would otherwise not be possible. This can be encouraged by incorporating the use of sonic expression within a variety of game design approaches; by addressing the primary component of game design, the in-game challenge, sound can be harnessed for the benefit of the overall playing experience.

Most games feature a challenge component for the player to overcome, requiring spatial reasoning to successfully navigate the game. Although there are a small number of games such as Metal Gear Solid (Konami, 1998), Hitman (Eidos, 2000), or Tom Clancy's Splinter Cell (Ubisoft, 2002) that use sound contextually to navigate an in-game challenge successfully, these challenges are ancillary to the main game challenge, with no significance placed on the quality of sound as expressed by the player. In these instances, there is a focus on minimising diegetic sound, as opposed to utilising sound for creative expression. The expression of the sound does not matter beyond the binary outcome (whether the sound is triggered or not). The use of music does appear in a small number of games where it is contextualised as a narrative device, inclusive of the game's main theme, such as The Legend of Zelda: Ocarina of Time (Nintendo, 1998), The Legend of Zelda: Majora's Mask (Nintendo, 2000), Eternal Sonata (Tri-Crescendo, 2007), Wandersong (Lobanov, G., 2018), and Brutal Legend (Double Fine Productions, 2019). However, in these instances, the resultant expression of sound is once again used as an event trigger as the player interacts with the game environment, solves puzzles, or performs attacks against computer-controlled characters. However, in these instances, the resultant expression of sound is once again used as an event trigger. However, there are a small number of games that contextualise sound as the primary in-game challenge component, such as Guitar-Hero (Harmonix, 2005). These are called 'music games' and have an in-game challenge relating to musical performance. Yet, players are not given the freedom to express themselves musically. Instead, these games reduce the performative musical experience down to a rigid sequence of instructions to follow. Once again, the expression of sound does not matter beyond the binary outcome, although in these games, it is vital for the completion of the in-game challenge. An even smaller subset of music games that are 'freeform' address this issue of creative freedom, but these games lack a clearly defined in-game challenge to solve. This practice-based research aims to address this gap, utilising sound as an expressive medium to create immersive and engaging playing experiences for games.

The term *playing experience* encompasses many factors, such as enjoyment, challenge, motivation, attention, and control. It also includes *immersion*. Broadly speaking, immersion comprises factors relating to perception and cognition. These can be framed as either sensory and psychological components (Carr et al., 2006), expectations of interaction (McMahan 2003), levels of attention (Cairns et al., 2014) (Jennett et al., 2008) (Grodal, 2003) (Klimmt, 2003), or modelled relative to: sensory experiences, challenge, and imagination (Ermi & Mäyrä, 2005). While this research is interested in the overall playing experience, the primary focus of the research is the measurable impact the use of sonic expression has on the player's sense of immersion. Through appropriate methodology, immersion can be analysed and evaluated, quantifying a player's level of engagement with a game (Cairns et al., 2014). The remaining playing experience factors are used to address the specific game design components implemented and evaluate their effectiveness in addition to supporting the findings related to immersion. This results in a comprehensive understanding of how different game components impact the overall playing experience. This process takes place over a series of three within-group studies that feature the experimental use of sound within game design. Each of these studies explores the relationship between sonic expression, immersion, and the overall playing experience of the game. Each consecutive study informs the next, and through this iterative design process, both game design practice and game design research methods are developed and refined in parallel. The findings of this research culminate into a cohesive conceptual framework for the use of sonic expression in game design, providing guidelines for any game designer to follow.

Providing a framework for others to follow is important. While there are resources available within the wider world of game design practice outside of academic game design research, these resources lack sufficient rigour to provide meaningful understanding. Game design approaches and techniques are frequently poorly documented, often informally presented, and seldom quantify how specific game mechanics impact the playing experience. Crucially, through the use of academic practice-based research methodology, this thesis will explore novel game components that utilise sonic expression and provide a measurable, repeatable outcome. This will result in a clearer understanding of the role for sonic expression in making immersive and engaging experiences. By concentrating these findings into simple guidelines to follow, this research provides new theoretical support for game designers. This is especially worthwhile for game designers who have little prior experience in sound design (Alves & Roque, 2011) (Collins, 2015). These guidelines may also prove to be illuminating for creative practitioners working outside of games, as other practices (such as interactive art) share commonalities in approach and outcome, intending to create highly immersive experiences. Additionally, outside of game design research, these findings may also prove useful for other research topics, such as broader Human-Computer Interaction (HCI), research reinforcement learning, and gamification. Beyond the contribution to the field of game design in academic research and the value of innovative approaches to follow in practice, this research also benefits my development as a creative practitioner as it refines my understanding of game design and strengthens my skill set.

Research Problem and Questions

The use of sound to immerse the player in games is often a passive response to the player's in-game interaction. Could the player's sense of immersion increase by offering the player a broader range of interaction methods for expressing themselves with sound? How does this alter the playing experience? By building games with novel sound interaction, we can study the impact increased sonic expression has on the playing experience. This will lead to novel game design approaches, sound design for games, methods generating procedural content relating to sound interactions, and provide insight into the impact player autonomy has on in-game challenge completion. To determine best practices and develop a framework for the use of sonic expression in game design, this research will address the following three research questions:

- To what extent does sonic expression impact the playing experience?
- Does increasing the depth of sonic expression produce a measurable change in player immersion?
- How does sonic expression impact game challenge, and does this affect immersion?

To understand how sonic expression can be used within game design, it is important to first identify salient factors of game design, along with the corresponding factors of the playing experience, that are influenced by sonic expression as a gaming mechanic. This can be achieved through testing appropriate platforms, experimenting within a variety of traditional gaming paradigms, evaluating the methods of control at our disposal, and designing novel in-game challenges. By answering 'To what extent does sonic expression impact the playing experience?' this research can highlight key areas of game design that should be explored further, helping to identify and prioritise factors of game design that will be the most effective in changing the player's experience. Likewise, establishing the relationship between sonic expression and player immersion is vital for our future understanding of how sonic expression can be used in games. If nuanced changes in sonic expression are reflected in the playing experience, then this will emphasise the importance of providing rigorous guidelines for game designers to follow. On the contrary, if there are clear perceptible limits (be they minimal or maximal) for the use of sonic expression, it would prove beneficial not only for informing design practice but also for technical optimisation. If the playing experience can be enhanced to the same extent using limited expression, then it would be unnecessary to offer a more complex method (regarding the minimisation of computational overhead from our software). Lastly, given that the primary novel component of most games is challenge (Brandse & Tomimatsu, 2013), it is imperative to understand the interrelation between challenge and sonic expression as measured in the player's sense of immersion. This information, backed by scientific rigour, would likely be a major incentive for any game designer to incorporate the use of sonic expression in their game.

By answering these questions through a series of three studies, this research will address the use of sonic expression within game design. This will determine the impact that sonic expression has on the overall playing experience, specifically on the player's sense of immersion. These findings will be formalised within a framework, providing guidelines for other game designers and future researchers to follow.

Objectives and Hypotheses

The overall objective of this thesis is to determine the impact that sonic expression has on the player experience and the player's sense of immersion. This research will also demonstrate novel gaming mechanics that make use of sonic expression and will identify areas where the use of sonic expression is suitable and recommend how it can be implemented. As there is little theoretical support for the use of interactive sound design within game design (Alves & Roque, 2011) (Collins, 2015), this research will be explorative. In addition to addressing the three research questions stated above, each study will concentrate on a different facet of game design that studies the impact of sonic expression.

The first study, 'Comparing the Playing Experience of VR and Mobile platforms when Using Sonic Expression as a Gaming Mechanic', has the specific objective of establishing whether there is any advantage to developing music games specifically for Virtual Reality (VR) over traditional screen-based platforms. The hypothesis is that there are differences in technological affordances for sonic interaction between screen-based platforms and VR, with

VR offering an advantage for the use of sonic expression to enhance the playing experience. Testing this hypothesis is important as it contributes to our understanding of interaction methods suitable for sonic expression and informs the remaining practical direction of this research.

The second study, 'Studying the Impact of Player Agency Over Rhythmic Sequencing and Platform Placement on the Playing Experience,' explores the impact of the player's creative decisions to navigate a typical in-game challenge. The aim is to determine the role of agency in music-game design and whether it can be leveraged to make engaging novel game experiences. The hypothesis is that offering players more agency in their musical-expressive decisions will lead to higher engagement and immersion. Examining this relationship between agency and the player's experience is vital for our understanding of what makes an engaging game challenge when using sonic expression, along with the impact that sonic expression has on the player's experience of the in-game challenge. It also establishes to what extent there is a dialogue between the player's musical expression decisions versus their motivation to complete an in-game challenge.

The third study, 'Music Game Design and the Impact of Non-diegetic Synchronised Cues on the Playing Experience', explores the importance of sound design decisions in relation to the player's interaction needed to complete the game challenge. It determines to what extent interactive sound design influences synchronised input. The hypothesis is that triggering non-diegetic sound, as opposed to diegetic sound increases the significance of the player's interaction, which impacts their playing experience. Challenging this hypothesis is important as it informs sound design practice and determines appropriate methods of interaction for specific types of sound. It also strengthens our understanding of how different types of sound can influence immersion in games. Answering these hypotheses and establishing the connection that the findings have in relation to the overall objective will help to answer the key research questions. In turn, addressing these research questions will help to identify the appropriate domains and necessary guidelines that will be featured in the framework that this research will develop.

Scope and Limitations

This thesis primarily investigates the impact of sonic expression on immersion and the player experience, but it may not account for all possible factors involved that influence these aspects of play. To limit factors outside of sound design, the use of sonic expression and the designed challenge component, the prototype games developed as part of this research removed or minimised any feature that was deemed unnecessary for the primary investigation. The in-game challenge is not contextualised within any wider narrative or story-telling element. There is no visual storytelling, and the art style and use of graphics are deliberately simple to not detract from the use of interactive sound. There is no competitive element to the games; each of the prototypes are single-player games with no score or leaderboard ranking their performance. When measuring the playing experience, we are not concerned with assessing aesthetic choices made within the sound design. Only in the context of sound interaction is the type of sound evaluated.

The research methods used do not explore all possible methods or techniques for implementing sonic expression within game design, as it primarily focuses on addressing the specific hypotheses outlined in the objectives. This thesis may not address every potential scenario or context in which sonic expression could be applied to game design. Game design is a broad, growing practice, and it is unfeasible to experiment with the use of sonic expression within every possible context. Likewise, the studies involve a specific set of platforms and game paradigms, and the findings may not necessarily generalise to all possible combinations of platforms and gaming mechanics. The scope of this research is limited to studying the impact of sonic expression in the context of game design technology and player experience; it does not delve into broader social, cultural, or economic aspects of game design.

The in-person studies were conducted within a controlled experimental setting, which may not fully replicate the complexity and variability of real-world gaming experiences. The scope of the in-person studies meant that immediate feedback could be given between researcher and participant. This was useful for issues relating to broader technology comprehension or specific technical difficulties. However, these were limited by reach, as participants were required to be physically present in a location. This potentially led to a less diverse sample due to participants availability to travel to a site in London. The remote studies had a broader reach of participants and thus could involve participants from a wider range of geographic locations, leading to a more diverse and representative sample. It also allowed for more flexibility in participant scheduling, making it easier to secure willing participants. However, a limitation of this study approach is that the researcher had less control over the participant's environment. Instructions could be given to the player, but it was difficult to confirm that they had been accurately followed. For example, in instructing the player to turn the volume of their device up to an appropriate volume level. Additionally, it was harder to ascertain the cause of any technical issues that arose during the remote studies. It was also evident that it was harder to maintain participant engagement and focus during the remote studies. Several participants for the third study (which was remote) did not play both games and did not complete both surveys. This resulted in their data being void for the study and removed.

It is essential to acknowledge that the results of this research may develop with advancements in technology and design practices, and therefore, the framework's relevance may change over time. The framework is not a definitive one-size-fits-all solution for the use of sonic expression in game design; but it does provide evidence-based guidelines to interpret and follow. Through further research and practice, these guidelines will evolve.

Research Methodology

This research is practice-based and explores novel approaches for sound design and game mechanics, investigating the relationship between sonic expression and immersion, bridging the gap between game design research and practice. Over three studies conducted within three different gaming paradigms - a puzzle game, a 3D platform game, and a rhythm game - these innovative approaches are evaluated to determine the impact that sonic expression has on the player's immersion and their overall playing experience. These studies employ A/B testing within groups and make use of a two-part survey given to the player after each playthrough. The game design approach and survey instruments are refined between each study. The research initially began with in-person studies but transitioned to include remote participation in response to the COVID-19 pandemic.

The games (A/B) developed for the first study were developed using C# in Unity (Unity Technologies, 2004 - present) and Visual Studio (Microsoft, 1997 – present) for flexibility across platforms. One version of the game was built for mobile (iOS/iPhone), while the other was built for VR (Oculus Rift DK (Oculus VR, 2014 - 2021)). For the second and third studies, both versions of the games (A/B) were programmed in JavaScript, using the A-Frame[1] framework (Marcos et al., 2015 - present) and MaxiInstruments[2] (Mimic Project, 2019 - present) library. These were built for in-browser play on a Quest 2 headset (Reality Labs, 2020). This switch of hardware was necessitated by the uncertainties of running in-person studies arising from the COVID-19 pandemic. This approach was chosen due to the popularity of the Quest 2 and the subsequent wide reach that the device would offer. Additionally, by delivering the study over a webpage, it provided low barriers to entry and ease of accessibility that would also benefit reach for study participation.

The surveys used within each study were comprised of two sections. The first was developed by this researcher and contained specific questions relating to game design elements of the prototype within that study. This section comprised binary, multiple-choice, and five-point Likert questions. Through an iterative development process, this section of the survey was refined to enhance participant engagement and efficiency while preserving data quality. The second part consisted of the Immersion Experience Questionnaire (IEQ) (Jennett et al., 2008). This was used as it is a validated questionnaire, and it provides valuable insight into the player's immersion as it subdivides the player's immersion into five immersion factors: 'cognitive involvement', 'emotional involvement', 'real-world dissociation', 'challenge', and

[1] A-Frame is a web framework for building virtual reality experiences. It provides an entity-component framework that is declarative, extensible and is composable to three.js (Marcos et al., 2015) (Cabello, 2010)

[2] MaxiInstruments is library featuring parameterised synthesisers and samplers that can be controlled in browser (Mimic Project, 2019 - present)

‘control’. This was useful for evaluating the player’s engagement relative to specific game design components.

Once anonymised and cleaned, data was analysed in SPSS Statistics. Internal reliability was checked for each part of the survey, and Cronbach’s alpha was reported, with improvements made when necessary. Data normality was assessed using Shapiro-Wilk tests due to the small sample size. Depending on normality assumptions, Dependent-means/Paired Sample T-tests or non-parametric Wilcoxon Signed-Ranks tests were employed for comparing means. Cohen’s effect size was calculated when applicable (measuring the effect size of the difference between two means). The results were analysed for the entire survey, along with each section separately. The results from the IEQ were examined across the five immersion factors. Demographic characteristics such as prior gaming and musical experience were correlated with playing experience and immersion. Evaluation of the findings generated additional game design questions and influenced the subsequent study depending on the feasibility and relevance of the findings. Potential areas of future research outside of the scope of this research were also suggested.

The prototypes developed integrated embodied sound and gameplay into the in-game challenge, prioritising these game components within game design. The games did not feature any narrative, and the game environment and art style were simplified. The in-game challenges were simplified into a base abstraction relevant to the specific gaming paradigm while incorporating mechanics to facilitate sonic expression. Consideration was made for each genre of the game featured in the research, as the different gaming paradigms have varying demands. For example, more importance is placed on the precision of sequencing in a rhythm game, as opposed to a 3D platform game where players can be afforded more time for their interactions. By specifically focusing on rhythmic interaction methods, a unified game design approach was maintained across the different paradigms. This allowed for iterative development across each prototype.

The sound design within each game was comprised of electronically generated samples and the use of synthesisers sequenced live in-game. Digital signal processing (DSP) techniques and effects were also utilised for player interaction. The prototypes developed for the third study addressed the use of both non-diegetic and diegetic sounds within the game. Decisions pertaining to musicality (pitch, timbre, harmony, tempo, instrumentation) and sonic aesthetic were made to align with the perceived game styles by this researcher. For example, the rhythm game adopted progressive techno as the genre, with a bpm of 127 (providing a manageable pace of the game), offering progressive complexity in note distribution while maintaining a minimalistic instrumentation of three synthesisers and a drum machine.

Organisation of Thesis

This thesis is divided into clearly defined chapters to effectively communicate the research undertaken. This introductory chapter primarily outlines the background, research questions, objectives, and scope and limitations of this research. While also summarising the research methodology, organisation, and significance and contribution of the research. The Literature Review chapter outlines the relevant existing literature and clearly defines all necessary terminology used within this research, such as '*What is sonic expression?'* It reviews the following topics of literature: sound, music and expression, existing sound-focused games, appropriate technical approaches suitable for the use of sonic expression in games, theoretical sound design approaches, presences and immersion, different models of immersion, flow, evaluating player experience, and metrics of evaluation player experience. The Methodology chapter outlines the broad approach taken within this research. As this research is practice-based, this chapter details the practical approaches in game design utilised in the prototypes for the three studies. This is broken down into the game design components of challenge, control, environment, art, and sound. This chapter also covers the survey design, how the research data was analysed and evaluated, and broadly outlines the technical delivery.

The following three chapters address the studies completed as part of this research. Each of these chapters contains the following relevant subsections: Abstract, Introduction, study specific Methodology, Game Design and Dynamics for the study, Combined Data Analysis for both sections of the survey, Data Analysis of section one, Data Analysis of section two, Data Analysis: Scoring the IEQ for each five immersion factors, breakdown of Demographics, Evaluation, Conclusion. All three study chapters aim to demonstrate novel game design practices. The first study chapter contains the Preliminary Study: Comparing the Playing Experience of VR and Mobile Platforms When Using Sonic Expression as a Gaming Mechanic. This studies the use of musical expression as the primary gaming mechanic. It explores the player's experience and sense of immersion across the two different gaming platforms. The outcome of this study determines what advantage there is to developing music games specifically for VR compared to traditional screen-based platforms.

The second study chapter contains the study of The Impact of Player Agency Over Rhythmic Sequencing and Platform Placement on the Playing Experience. This study explores the impact of the player's creative sound decisions to navigate the typical 3D platformer game challenge on their enjoyment and sense of immersion. The findings of this study help to elucidate the role of agency in music-game design and explore the relationship between musical expression and challenge. The third study chapter explores Music Game Design and the Impact of Non-Diegetic Synchronised Cues on the Playing Experience. This study explores the impact of sound design on the playing experience and sense of immersion. It aims to ascertain to what extent interactive sound design choices can encourage synchronised input. The study aims to determine how player interaction can be used in sound design and whether increasing the significance of the sonic outcome derived from the player's interaction has an impact on their playing experience. The findings from these studies are then discussed in the following chapter.

The Discussion Chapter contains a full breakdown of all results produced by the three studies. The research questions and objectives are reaffirmed, and a summary of the findings is presented. The findings are then interpreted and analysed, explaining what the data means in context to the three key research questions of this thesis: To what extent does sonic expression impact the playing experience? Does increasing the depth of sonic expression produce measurable changes in player immersion? How does sonic expression impact game challenge, and does this affect immersion? The outcome leads to the basis for a conceptual framework for the use of sonic expression in game design. This is the DIVE framework, and it is split into four connected domains: Demand, Inclusivity, Versatility, and Engagement. The remaining chapter contains discussions on the Contribution of the research, Reflection on

Methodology, Limitations of the research, Practical Implications and Future Research, and a conclusion of the discussion.

The last chapter is the Conclusion. This chapter restates the research objectives of the thesis and summarises the key findings, highlighting the most significant results. These findings are used to address the original research questions. The constraints of the findings are discussed, along with the strengths and limitations of the methodology. The contributions to the field of game design are covered, additionally discussing the broader implications of the research for the field and beyond. Suggestions for future directions that could be built from this work are offered. Lastly, the conclusion includes a brief personal reflection on this research.

This thesis also contains a comprehensive list of all references used and an Appendix. The Appendix contains the game design documents for each of the three prototypes. It also contains copies of the preliminary pre-survey questions, section one questions, section two questions, IEQ questions used, and a script of the instructions for the first study. There are also copies of post-surveys for the second and third studies.

Significance and Contribution

This practice-based research makes a dual contribution, primarily focusing on innovative sound design approaches and novel game mechanics. Each of the prototypes developed for the three studies are unique, featuring original gameplay elements that differentiate them from existing games. By removing the use of strict input-matching and making use of real-time sound generation, sonic expression can be used to navigate in-game challenges, creating engaging player experiences, and positively impacting their immersion. Additionally, the research introduces the DIVE framework. The framework, developed through player experience surveys developed by this researcher and the use of the validated IEQ, provides practical insights for sound designers and game developers. Providing a framework with guidelines is an important contribution, as there is little in the way of theoretical support for game designers, especially those inexperienced in sound design (Alves & Roque, 2011) (Collins, 2015).

The research also has implications for the broader game design community, as it offers tangible examples for designing engaging experiences with sonic expression. Researchers in more general human-computer interaction fields may also benefit from the research, along with those working with reinforcement learning and gamification. Practitioners from other creative fields, such as computational arts or interactive design, may also find the research valuable, as some of the techniques are directly applicable to their practice. Additionally, this research opens further opportunities to explore the impact of sound on the playing experience. For example, to what extent does previous experience impact cognition and challenge in immersion? What is the relationship between dynamics, static musicality, and immersion? Or What is the significance of continuity in immersion? Studies in these areas of enquiry (or others pertaining to sonic expression in game design) can interrogate and expand the DIVE framework, validating it as a useful resource for game designers.

Conclusion

Sound is often used passively in game design to enhance the playing experience. When there is interactive sound in games, it usually acts as a binary event trigger. Even in games where sound is contextualised within the story, there is no inherent value in the player expressing themselves with sound. In games where the challenge is entirely contextualised in music, there is limited expressivity, which is strictly enforced by the rigidity of the in-game challenge. By studying comparable creative electronic music practice, we can identify approaches to electronic music interface design that can be used within game design. The application of this would be novel and highly innovative within the practice of game design. The impact of these unique gaming mechanics will be evaluated by measuring the playing experience, specifically evaluating the player's sense of immersion. This will take place over a series of three separate studies.

For each of these three studies, two versions of a novel game prototype are developed. Each study contains a two-part survey. The first section is a survey on player experience devised by this researcher. The second section of the survey contains the validated Immersive

Experience Questionnaire (IEQ). This provides a breakdown of the five immersion factors of 'challenge involvement', 'emotional involvement', 'real-world dissociation', 'challenge', and 'control'. This questionnaire provides a firm understanding of the player's engagement. The surveys are used to test a separate hypothesis within each study. These relate to the role of sonic expression, agency, sound design, in-game challenge, and overall game design. The overall objective of this research is to determine the impact sonic expression has on the playing experience and the player's immersion. The findings of these studies will help answer the three key research questions of: To what extent does sonic expression impact the playing experience? Does increasing the depth of sonic expression produce a measurable change in player immersion? How does sonic expression impact game challenge, and does this affect immersion? Through appropriate analysis and evaluation, by addressing these questions with our findings, this research will provide the conceptual (DIVE) framework that will help support other game designers in implementing the use of sonic expression with game design practice.

This chapter has also broadly outlined the methodology of this research and highlighted the scope and limitations of the approaches used. This includes identifying what game design components were not included within the prototypes, the limitations of the included components, and identifying the scope and limitations of the study design. The organisation of the thesis has also been communicated, highlighting what is in each chapter. Lastly, the significance and contribution of this thesis has been broadly outlined. As this is practice-based research, the contribution is split between game design research and games game design practice. The practical contribution is the novel gaming mechanics and innovative sound design approaches that were developed. The research contribution is codified as the conceptual DIVE framework, serving as vital support for other game designers. The research has implications for related creative practice and other fields of research.

Sonic Expression, Immersion, and the Playing Experience: A Review of Current Literature

Introduction

This research seeks to understand how player experience is impacted by the player's sonic expression. To explore this question, research will be conducted through the development of video-game prototypes (hereon referred to as games or prototypes) in which players use different forms of sonic expression to interact within the game environment. These methods are derived from, or are otherwise similar to, those found in sound design, electronic music, and related audio-visual creative practice. This is because these approaches offer well-understood mechanisms for sonic interaction and expression and are discussed later in this chapter under the heading of 'What is Sonic Expression?' By coupling player interactions to sound components and using these components to generate the challenges faced within the game, the aim is to explore the impact that interactive sound has on player immersion. In doing so, this will demonstrate novel sound-design techniques specifically for games. A series of prototypes will be made, presenting the same core gaming mechanic. This mechanic requires the player to generate sequences of sonic events of their choosing to solve game challenges presented in different forms. These prototypes will examine the mechanic within three gaming paradigms: a puzzle game, a rhythm game, and a platform game. Each prototype will allow the player to impact the sound to a different extent.

To find the most effective methods to answer the research questions laid out on page 20 of the introduction chapter, we must first review existing material from related research fields and define our terms concerning each fundamental component of the research questions. What constitutes sonic expression? What is the state of the art with respect to sonic expression through digital technology? What existing games offer expressive or immersive sound experiences? What are the existing challenges and approaches to game sound design? How has immersion been studied in games?

What is Sonic Expression?

Before reviewing sound design in games, the related player interaction, and the resulting experience, it is important to define what we mean by sonic expression. We can summarise sonic expression to mean any of the player's actions that have an impact on the sound of the game and over which they feel they have sufficient control. The practical output of this research will be centred around sound and music interaction in games, and as our games are digital, we can also include editing, sequencing, and Digital signal processing (DSP) techniques (a simple example being filtering) that have now become commonplace in electronic music production and sound design. Subsequently, 'sonic' is the appropriate term as the player may not necessarily be manipulating the conventional musical content (such as tones, harmonic quality, and rhythm intervals) but the soundwaves within the game. Expression, on the other hand, is harder to define, as even within music, it appears in many different contexts.

One can reduce musical expression to 'the act of communicating' (Fels et al., 2002) that takes place both in the player and the listener's perception of the performer's actions in relation to their sensory experience. Therefore, expression can describe the organisation of sound elements, such as specific note intervals and durations (Dobrian & Koppelman, 2006). It is also used to describe the stress or articulation of specific notes (for example, vibrato on a violin) (Fels et al., 2002). In performance, it can be considered the realisation of these compositional elements, as well as the applied interpretation of the composed material through the use of gestures and other techniques of the player (Dobrian & Koppelman, 2006) (Tanaka, 2010).

As this research is focused on the player's experience, or in musical terms, our performer's experience, the listener's perspective and experience fall outside of the remit of our research. Subsequently, I will only need to consider 'the transparency of device mapping' (Fels et al. 2002) for the performer. Mapping, in the context of technological affordances (Gibson, 1979) (Gaver, 1991), is important for situating the player's input gesture to the sound and corresponding game function (Tanaka, 2010) (Poepel, 2005). Successful mapping distinguishes how specific input can be used, resulting in an intuitive interaction for the player. This is favourable for teaching players how to be expressive within the software.

(Tanaka, 2010). It is also a factor when evaluating the role of specific sounds within the game and how they can affect the player (Pichlmair & Kayali, 2007) (Kurtz, 1998). Pichlmair & Kayali found through qualitative analysis of seven different qualities of music games: active scores, rhythm action, quantisation, synaesthesia, play as performance, free-form play, and sound agents, that there were two separate types of music game (rhythm games and instrument games) (Pichlmair & Kayali, 2007). While this research is not about building new interfaces, the importance of mapping player expectations for their interactions will be reinforced later when we discuss conditions required for game immersion. This is because the player's expression will be used to control a musical system, with the sense of immersion arising due to the feeling of control that players may have over sound interactions. The impact of the related technological limitations on the design of the game will also be important to consider, as they could also impact the player's immersion.

Research in new interfaces for musical expression can carry with it connotations of 'high-art' or the demands of a classically situated musician, often equating the range of 'virtuosity' (Dobrian & Koppelman, 2006) to expressivity. However, it is still relevant to my domain as the in-game challenges require the players to perform with sound in a capacity similar to that of a musician, albeit whilst also completing in-game tasks. Additionally, even if a device has a comparatively limited number of axes to be expressive in, it does not necessarily equate to the object being any less expressive. This is because to be expressive, as defined by Dobrian & Koppelman, is to 'effectively conveying meaning or feeling' (Dobrian & Koppelman, 2006); if it is effective at communicating, then it is expressive. A clave is still capable of expressing a range of qualities, especially in the hands of an expert player. Although offering limited range in the form of pitch or timbre, a clave player still makes use of intensity (how hard the clave is hit to accent the beat) and timing. As noted by Iyer in studying the expressivity of micro-timing in African-American music,

> 'Miniscule timing deviations in performed music are often misleadingly described as "discrepancies" (Keil & Feld, 1994) …It turns out that these deviations from strict metronomicity both convey information about musical structure and provide a window onto internal cognitive representations of music' (Iyer, 2002).

In rhythmic or 'groove' based music ('music that gives rise to the perception of a steady pulse' (Iyer, 2002)), performers are able to,

> 'Evoke a variety of rhythmic qualities, accents, or emotional moods by playing notes slightly late or early relative to a theoretical metric time point. Numerous studies have dissected the nuances of expressive ritardandi and other tempo-modulating rhythmic phenomena (Desain & Honing, 1996) (Friberg & Sundberg, 1999) (Repp, 1990) (Todd, 1989), whereas fewer careful quantitative studies have focused on expressive timing with respect to an isochronous pulse (Bilmes, 1993) (Collier & Collier, 1996) In groove-based contexts, even as the tempo remains constant, fine-scale rhythmic delivery becomes just as important a parameter as, say, tone, pitch, or loudness.' (Iyer, 2002)

In order to navigate the space between gesture, or in our case, the much broader term, interaction and sound, one must address the process of how humans engage with music (Leman, 2007). However, while interesting, the mediation of the whole music experience will not be explored, as the scope is far too broad. Instead, we will confine corporeal music engagement to the small, intended actions, the corresponding interaction, the consequential descriptors (Wanderley & Orio, 2002), and the accompanying in-game challenge. Corporeal music engagement pertains to the embodied experience and the physical actions involved. The in-game challenge will situate the embodied sound and game action together. Meaning the act of expressing sound is the game action required by the game challenge.

Game Challenge

The core component of most games is the 'challenge', an overall objective or goal for the player to achieve (Jamieson, 2017). This is explained in more detail by Brandse and Tomimatsu, who state,

> 'To games, challenge is a very important element to create a good user experience (Cox et al., 2012). This is further confirmed by Johannes Huizinga, who stated that games are largely about overcoming something (Huizinga, 1968), giving further weight

> to the notion that challenge is integral to games. Juul, J., also heavily hints at this in his definition of what is a game, as games need to be “challenging” (Juul, 2003)’ (Brandse & Tomimatsu, 2013).

As identified by Brandse and Timimsatu, both Huizinga and Juul allude to the player being responsible for achievement in the context of playing a game. This is further supported by Juul’s study of players’ fear of failure:

> ‘Players clearly prefer feeling responsible for failing in a game; not feeling responsible is tied to a negative perception of a game. In effect, this sharpens the contradiction between players wanting to win and players wanting games to be challenging: failing, and feeling responsible for failing, makes players enjoy a game *more*, not less.’ (Juul, 2009).

This perception of responsibility from the player closely relates to the concept of ‘non-trivial interactions’ (McMahan, 2003) and the concept of ‘discernible outcomes’ (Sale & Zimmerman, 2003) as covered by Ermi and Mäyrä (Ermi & Mäyrä, 2005) and pertain to immersion, which is covered later in this chapter. Players must feel challenged, and their actions must have consequences for this challenge to be meaningful.

Another important factor of game challenge to consider, as highlighted by Cox’s work measuring the game experience is that ‘…the definition of challenge is purely in terms of players’ perception of difficulty.’ (Cox et al., 2012). It is dependent on the player; the challenge may be informed by one aspect of the game’s design as opposed to a different aspect for another player. Therefore, it is important to know how the challenge is framed. Cox highlights that typically:

> ‘There are two main ways to achieve this: push the gamer’s physical limits; or push the gamer’s cognitive limits (Hsu et al., 2007) (Orvis et al., 2008). The gamer’s physicality limits the speed with which interactions with the game can be conducted, and the accuracy with which actions can be performed. The gamer’s cognitive abilities have a

> limiting effect on speed and accuracy of the problem-solving activities required by the game.' (Cox et al., 2012).

This divides the game challenge into two distinct domains. However, it could be argued that any game challenge is a combination of these two domains; even if a challenge is heavily skewed towards physical limits, it will likely still have some cognitive demand and vice versa. The overall cognitive and physical limit forms the challenge for the player. Considering these limits relative to the use of sonic expression will be important when designing the prototypes, as it is likely a challenge featuring a high level of sonic interaction will be novel for the player.

Music Games

Music games is a term often used interchangeably with the term *rhythm games*. They predominantly hinge their entire gameplay around sonic interactions, although, as will be explained, the audio is not often interactive. These types of games are not to be confused with audio games, in which the entire game environment is set in audio only, such as Papa Sangre (SomethingElse, 2013) and Real Sound: Kaze No Regret (WARP, 1997); audio games are not within the remit of this study. Rhythm games are a small but highly profitable market of games, often sold with plastic peripherals that mimic an instrument's shape but with a simpler interface. Rhythm games often use popular music as either a context for generating a pattern in which the player must input accordingly in time with the music (Guitar-Hero (Harmonix, 2005), Rock-Band (Harmonix, 2008), Parappa the Rapper (NanaOn-Sha, 1996)), or where certain qualities of the audio file of the music itself procedurally generate the patterns (Vib-Ribbon (NanaOn-Sha, 1999), Audiosurf (Fitterer, 2008)). Pichlmair & Kayali state, 'rhythm games offer little freedom of expression apart from the prerogative to perform while playing. They strictly force rules on the player on how she has to react to a specific stimulus displayed on screen or communicated by sound.' (Pichlmair & Kayali, 2007)

According to Pichlmair & Kayali, these strict rules mean that despite rhythmic sequence matching being performative (and subsequently utilising player expression in some capacity), there is very limited scope concerning how the player can be expressive with sound in rhythm games. Any independent sonic expression is punished for deviating from the intended outcome (Pichlmair & Kayali, 2007). While the player's imagination may bridge the gaps between the complexity of the task and the musical output (Ermi & Mäyrä, 2005) (Grimshaw, 2008) (a concept we will explore further in this chapter under the heading of 'sensory and imaginative immersion'), the fact the interaction is limited to a desired response to predetermined series of events could impact the playing experience. As it only limits the player's sonic expression to the game designer's intended outcome, and, in turn, this results in a very similar experience of the game environment in each playthrough.

Pichlmair & Kayali are not as severe when defining what constitutes a device for musical use, stating that although mastering an instrument imposes its own challenge, 'many toys are played in an explorative way and instruments can be played in that way, too. According to Kim (Kim, 2004), games are mere toys plus challenges and goals.' (Pichlmair & Kayali, 2007) This alludes to an intersection of what is an instrument, what is a toy, and what is a game. So, rather than an innate quality of the device that makes it musical, it is rather the approach of the player or framing of its use. According to Kim, 'games are rule-based systems in which the goal is for one player to win.' (Kim, 2004). Where human opponents are absent, the goal is then against the challenges imposed by the game designers. Without a goal and the corresponding challenges that go with it, games would be a toy. This may be a simplistic view.

Freeform Play

Rhythm games are not the only music games, and there have been a few that take almost a completely opposite approach in their design – offering 'freeform' gaming experiences. Unlike rhythm games, freeform games tend not to have an explicit game challenge for the player to complete. Instead, they offer the player a game environment for the player to explore, usually at their own pace. The player is 'free' to 'form' their own use out of the game. Whilst freeform music games tend to afford the player more opportunity to be expressive, it is often at the expense of the gaming challenge. The flagship PS4 title for Sony's new virtual reality headset – Harmonix Music VR (Harmonix, 2016), although impressive, seemingly lacks the most basic attribute of what constitutes a game – the challenge component. To a lesser extent, the same could be said for Nintendo's Electroplankton (Indieszero, 2005). Although Electroplankton offered a range of interactive sounds for the player to utilise, it only had a small game challenge, which served more of an instruction guide for the freeform mode than a challenge for the player to be engrossed with.

The Playstation title Music 2000 (Jester, 1999) is even more free in its approach, as it completely removed the challenge component. Despite being released on a gaming console, it was not, in fact, a game but sequencing software. As discussed in this chapter above (page 38), instruments (or, in this case, electronic music production tools) or toys can all be used in an exploratory way as a method of expression. However, despite not being a game, the importance of this title should not be overlooked. By being more akin to digital audio software like Cubase (Steinberg, 1989-present) or Logic Pro (Apple, 2002 - present, previously CLAB Notator) than a game, it introduced electronic music-making concepts to many people who otherwise may not have encountered them (Baines, 2015). These tools can potentially be very expressive, providing the user invests their time accordingly. This has the impact of expanding the influence of video game music on contemporary music production.

The Dreamcast title Rez (United Game Artists, 2001) (a 'remastered and evolved' sequel Rez Infinite (Monstars, 2015) is also scheduled for release on Sony's VR in 2016) has game challenges but also allows the player some degree of autonomy over how their game interactions overlap their sound interactions. Arguably, the player's sound interactions are passive and entirely dependent on the required interaction needed to successfully navigate the gaming challenge. The game is not designed around the idea of the player expressing themselves with sound. Rather, the game encourages an optimal harmonious interaction between the spatial navigation of the level and the sounds of the game. This method of overlapping sonic interactions with the required challenge interactions is also apparent in the Playstation VR (and later other platforms) title Thumper (Drool, 2016). Overall, there seems to be an emerging dichotomy between music games. Either they have a clear game challenge but offer very limited features in expressivity, or they include a variety of different tools affording the player a broad range of expression but lack a substantial game challenge.

Procedural Content Generation in Music Games

Although there are many ways in which game content can be created, the most relevant development to this research is procedural content generation (Hendrikx et al., 2012) (Risi et al., 2014). Using procedural content generation to generate environmental content from the player's expression and corresponding interaction will present them with a novel playing experience each time. If the consequences are non-trivial in the game environment, the results should impact the player's sense of immersion (Ermi & Mäyrä, 2005) as outlined later in this chapter on page 47. The term 'non-trivial' is important to note, as Connor concluded, 'there is sufficient evidence that suggests that there are particular areas where the procedurally generated content does not succeed in engaging the player as well as human-designed levels.' (Connor et al., 2017). This was tested by measuring player experience (Jennet et al., 2008) across two different versions of a 2D top-down shooting game, where one version had human designed levels and the other version featured procedurally generated levels. The hypothesis was that immersion ratings between the two variations would not be significantly different. Although the results were found to be statistically inconclusive, cross-referencing specific questions of the player experience questionnaire suggested that human design game levels had achieved a higher level of immersion. (Conner et al., 2017). This suggests that it would be important to consider what particular combination of player expression and level design one chooses to use procedural content generation for, as it may potentially have the opposite effect. However, if this approach was applied to the in-game challenges, it may force the player to be more attentive. The notion of challenge modelling for procedural level creation has been covered by both Sorenson (Sorenson et al., 2011) and Smith (Smith et al., 2011). Both build on the concept of 'rhythm groups', which take the difficulty of in-game tasks, divide them into high and low challenges and arrange the challenges according to a ruleset. Smith specifically focuses on grammar for parameters that influence level pacing and geometric design. Given that all musical events exist at a base level of periodicity, it would be interesting to see a similar approach to Smith's being used to generate or modulate a game environment based on the player's musical input. Smith also points out that while there is a small amount of literature on game design, it is primarily regarding 'macro' gaming elements. 'Short descriptions of genre-specific level design are not sufficient to provide a detailed understanding of the structural interrelationships of elements in a level.' (Smith et al., 2011)

Sound Design in Games

Despite sound being a valuable component of the overall game aesthetics and affective perception (Lennart et al., 2011), there is little theoretical support for game designers, particularly for those inexperienced in sound design (Alves & Roque, 2011) (Collins, 2015). Alves & Roque argue that this is in part due to the observation that HCI research in computer games is often 'directed towards visual modality, leaving others, like sound, less explored' (Alves & Roque, 2009). This is also stated in Chapter 17 of Game Sound Technology and Player Interaction: Concept and Developments (Alves & Roque, 2011), where Alves & Roque state, 'practices on game sound are strongly influenced by those from cinema.' (Alves & Roque, 2011). NG & Nesbitt argue this visual-orientated focus is detrimental to the development of sound in games. They state that is 'a consequence of this trend towards designing visually enhanced and realistic playing environments is that the design of sounds in games is often neglected or only used to decorate the visual design.' (NG & Nesbitt, 2013). This sentiment is shared by Kenwright who states, 'Often audio is an afterthought in many interactive environment projects, with the assumption that sound is not the reason the project fails.' (Kenwright, 2020). Although Kenwright does go on to say 'Sound (like all aspects) should be designed, whether intentionally or not. In most cases, bad sound design is worse than no sound design at all.' (Kenwright, 2020) which highlights the important sound can on the player's experience. This is supported by Cunningham (Cunningham et al., 2006). After surveying gamers regarding what features they considered important when purchasing games and finding that playability ranked the highest with sound the lowest, Cunningham noted, 'It is probably that a lot of developers chose not to risk large sums of money on new ideas for audio technology, only to find that it only appeals to a small audience.' (Cunningham et al., 2006). However, it is also important to note that whilst 'playability' ranked as the most important factor, sound does, in fact, play an important role in just how playable a game is (Tan, 2014). Tan's findings are further supported by the later research by Wöhrman & Ningalie titled 'The Impact of Sound Player Experience', where they found 'some parts of the game world might be ignored by players as a result of poorly integrated background music or sound effects.' (Wöhrman & Ningalie, 2018). This was determined through an increasing in heartrate of the participants or an observed change in the participants playstyle. Additionally, in related research by Tafalla, gamers have been shown to perform better with in-game sound turned on (Tafalla, 2007). This is further supported by Cassidy & MacDonald who found that self-selected music improved participant performance and their affective state (Cassidy & MacDonald, 2009). The notion that players do not purchase a game based on the technical qualities of the game is referred to as 'the performance trap' by Wesley & Barczak:

> 'Designers and engineers are often energised by breakthrough technologies that allow them to accomplish tasks they only dreamed were possible. In the process, they often lose sight of the real goal – fulfilling a customer need. They succumb to what we call "the performance trap."' (Wesley & Barczak, 2010).

You could argue that this occurs when game manufacturing is driven by marketing, as players do not purchase games for their love of programming.

The game's genre is also a factor in how important or in what way the sound can influence the player's experience or affect the playability of the game. Yamada studied the impact of music in a driving game, concluding that 'the negative effect on performance may be a result of music disturbing the players' concentration.' (Yamada et al., 2001) (Yamada, 2002) This conclusion is based on the possibility that there was a complex relationship affecting the playability between the audio and visual components of the game. This is further explored in Lipscomb & Zehnder's later study on the impact of a musical score on the gaming experience (Lipscomb & Zehnder, 2005). Unlike Yamada, Lipscomb & Zehnder used a role-playing-game to test the effects of a musical score on player immersion, proposing that,

> 'The musical excerpts used by Yamada were selected from a wide range of varying musical styles…the results of the study are likely to be significantly different from those in which a musical score has been specifically composed to accompany events as a dramatic narrative unfolds.' (Lipscomb & Zehnder, 2005).

However, despite differences between the two studies, Lipscomb & Zehnder found a similar complex relationship between audio and visual components that can affect playability; some players' responses included transference between audio and visual components, meaning the nature of the audio impacted the player's experience of the visuals, and vice versa. This is further supported in a later study on the audio influence on game atmosphere by Andersen, who 'found out that audio doesn't only affect the gameplay of the player but also the satisfaction factor of doing an action, based on the feedback of the participant of our game test.' (Andersen et al., 2020).

Existing theory on the subject of sound in games maintains the distinction between diegetic and non-diegetic sounds. Originating from the study of literature, it is also used in film studies (Chion & Gorbman, 1994). Diegetic sounds are used to situate the player within the game world, such as the sound of a weapon being drawn. Whereas non-diegetic sounds often support the narrative, such as a musical score or the user interface emitting an alarm when running low on health to alert the player (Russell, 2012). Whilst there are parallels or commonalities between games and film, there are also vast differences between the two. Collins points out that,

> 'a composer of music for linear media can predict how the music will sound from beginning to end for the listener, and compositions are constructed with this aspect taken for granted. The music of non-linear media such as video games, however, works more like a major urban metro: At any time, we may want to be able to hop off at one station and hop onto another train going in a new direction.' (Collins, 2007).

In addressing the non-linear nature, Collins encompasses both 'interactive 'and 'adaptive' audio under the term 'dynamic audio', with interactive audio pertaining to sounds directly related to the player's input, and adaptive audio relating to how sound and music change appropriately in relation to gameplay. Both can be used for diegetic and non-diegetic sounds, although interactions tend to be diegetic, and adaptive tends towards non-diegetic sounds.

A supporting model that builds on the traditional two-dimensional diegetic/non-diegetic framework is the IEZA model presented by Huilberts & Van Tol (Huiberts & van Tol, 2008).

The IEZA model relates these two conceptual dimensions to what is communicated by the in-game sound, offering four domains of 'Effect, Zone, Interface, and Affect.' *Effect* relates to the player's activity in the diegetic part of the game. *Zone* relates to those linked to the in-game environment. *Interface* is used to express activity in the non-diegetic part of the game environment. Finally, *Affect* is linked to the non-diegetic part of the game environment that communicates the narrative or 'setting' of the game. Whilst this often features music, it is worth noting that this can sometimes be carried over from the diegetic game environment.

Huiberts states that game audio and the IEZA is primarily concerned with two different functions. The first is to optimise gameplay (as expressed by Cunningham regarding the support for playability), as well as to 'dynamise gameplay', meaning to make the gameplay experience more engaging (Huiberts, 2010). The latter is suggested by Collins as a demonstration of how successful designers juggle dynamic audio components to impact the user experience. 'Moving smoothly from one music cue to another assists in the continuity of a game, and the illusions of gameplay, since a disjointed score generally leads to a disjointed playing experience, and the game may lose some of its immersive quality.' (Collins, 2007).

Whilst the practical output of this research focuses on music games, specifically ones that require the player to make sonic expressions, they should fit within the game as part of a cohesive whole. Design conventions that impact how programmers decide how sounds change, including methods for avoiding listener fatigue, will be important. (Collins, 2007) In 'An Introduction to the Participatory and Non-Linear Aspects of Video Games Audio,' Collins discusses some of these methods: 'composers now commonly re-use cues in other areas of a game, to reduce the amount of unique cues needed, but without creating a repetitive sounding score. This requires careful compositional planning, and often a reduction in dramatic melody lines so that the cue is less memorable.' (Collins, 2007). another approach is the use of incorporating 'timings into the cues, so that if the player does get stuck on a level, the music will not loop endlessly, but will instead fade out.' (Collins, 2007). Finding the appropriate design solution to listener fatigue is important as we want players to be encouraged to make sound and express themselves, not avoid it.

A more recent development in sound design for games has been the use of procedural audio, where sound is made algorithmically by using generative models. This sound design is not pre-recorded or pre-sequenced but determined by game data or in-game events. Yee-King & Dall'Avanzi define the term generative 'to refer to models utilising algorithms to produce an output that is not explicitly defined. A simple example is writing several fragments of a musical score, then choosing the sequence in which the fragments are played back at random. This 'musical dice game' generates a different piece of music each time it runs.' (Yee-King & Dall'Avanzi, 2018). We can find examples of procedural audio for music generation in Rise of The Tomb Raider (Square Enix, 2015); where it was used to generate percussive sequences in the soundtrack, and for sound design in No Man's Sky (Hello Games, 2015); where it was used to generate the voices of procedurally generated animal characters. While this approach addresses some of the challenges faced by sound designers as described by Collins above, it does come with its own new challenges. Yee-King & Dall'Avanzi specifically highlight: 'Procedural audio faces several challenges to its broader adoption - it demands a different and more extensive skillset than fixed asset production, it can be difficult to match the sound quality of fixed assets, and it has a different computational resource profile.' (Yee-King & Dall'Avanzi, 2018).

Presence and Immersion

Although presence will be outside the remit of this research, it is important to consider that the study of presence encompasses, makes use of, and propagates study into immersion. Presence is the sensation of *'being there'*, - and incorporates the notion of virtual embodiment, whilst immersion is defining the technical terms of how presence can occur (Grimshaw et al., 2011) (Reiner & Hecht, 2009) (Slater et al., 2009). It has become more significant since the release of the HTC Vive, Oculus Rift, and Sony PS4 VR virtual reality (VR) headsets - the next generation of gaming peripherals. As discussed by Grimshaw & Charlton, there are two opposing arguments for what presence is. Reiner & Hecht state presence takes place in the physical position within a virtual environment, whereas Fencott and Slater argue that presence is cognitively experienced by the user (Fencott, 1999) (Slater, 2002). Both arguments infer that immersion depends on environmental factors of the game impacting the user. Slater states,

> ‘We reserve the term immersion to stand simply for what the technology delivers from an objective point of view. The more that a system delivers displays (in all sensory modalities) and tracking that preserves fidelity in relation to their equivalent real-world sensory modalities, the more that it may be described as being “immersive”’ (Slater et al., 2009).

We will take the ‘virtual environment’ in the instance of games to mean the game environment, regardless of how the player is represented or ‘virtually embodied’ within this space (be it with a 3D model, 2D sprite or a conceptual framework of non-visual interaction, i.e., Tetris). Subsequently, using Slater’s definition, everything within the game can help immerse the player. Cairns uses this idea to build a hierarchy of immersion, initially defining immersion in the following way: ‘immersion… is intended to mean the engagement or involvement a person feels as a result of playing a digital game.’ (Cairns et al., 2014). This is in stark contrast to Slater’s view that immersion is what is technologically deliverable, as mentioned above. Rather, it relates more closely to Slater’s ‘Notion of immersion as part of a potential understanding of the meaning of *presence*: Immersion is a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences.’ (Slater, 1999) Cairn’s model of immersion was expanded when Brown and Cairns conducted a qualitative, grounded theory study with players, finding that:

> ‘Players were able to distinguish different levels of immersion in games and these corresponded to their sense of engagement and involvement in the game. The basic level of immersion is *engagement,* where players simply invest time and effort to play the game. The next level is *engrossment,* where players are dedicating attention and emotional involvement in the game. The third and highest level is *total immersion* which they identify with presence.’ (Cairns et al., 2014)

Brown and Cairns describe ‘attention’ as ‘a willingness to concentrate’ (Brown & Cairns, 2004). Whilst Cairns’ use of presence to define total immersion is perhaps confusing, they attempt to clarify as follows: ‘Total immersion was seen as the idea of complete involvement with the game where nothing else matters and the player feels "in the game."' (Cairns et al., 2014).

Sensory and Imaginative Immersion

Cairns' cognitive basis of immersion works in conjunction with a sensory basis, which, as Grimshaw notes (Grimshaw, 2008), works similarly to Ermi and Mäyrä's sensory, challenge-based, and imaginative immersion model (SCI model). In Ermi and Mäyrä's model, sensory immersion relates to the audio-visual 'physical' components of the game environment, challenge-based immersion relates to the use of motor or mental skills, and imaginative immersion covers narrative and affective components. Unlike Cairns' hierarchy of immersive states, the SCI model surmises immersion to be varying levels of all three components at the same time (Ermi & Mäyrä, 2005). This allows for the superposition of the two theories; for example, in relation to sensory immersion, 'the audio-visual, functional and structural playability… can be seen as prerequisites for gameplay immersion and rewarding gameplay experiences' (Ermi & Mäyrä, 2005), implying that both sensorial components and the challenges they can provide will facilitate engagement. Additionally, if either of these is sufficient, the player could be considered 'dedicating attention' (Jennett et al., 2008) (Grodal, 2003) (Klimmt, 2003), which would be halfway to engrossment (Brown & Cairns, 2004). The remaining 'emotional involvement' would then be comprised of the 'imaginative immersion' – whether or not the player can form a contextual understanding of the surrounding narrative of the challenge. Lastly, total immersion would simply be an optimised SCI model – where the constitution components show an average high enough for the player to be totally immersed.

Similarly, other theorists make the same distinction between sensory and psychological components of immersion. Carr splits immersion between sensory and psychological factors (Carr et al., 2006), whereas McMahan divides immersion into three conditions hinged on the player's interactions,

> '(1) the user's expectations of the game or environment must match the environment's conventions fairly closely; (2) the user's actions must have a non-trivial

> impact on the environment; and (3) the convention of the world must be consistent, even if they don't match those of "meatspace."' (McMahan, 2003)

All three require both the sensorial and cognitive faculties of the player. Considering my focus is oriented around player interaction and the affordances players are offered, McMahan's notion that interactions must meet expectations for immersion to occur seems like the most relevant. In conjunction, Ermi and Mäyrä relate McMahan's 'non-trivial impact on the environment' to Salen and Zimmerman's (Salen & Zimmerman, 2003) definition of meaningful play, stating that it:

> 'Occurs when the relationships between actions and outcomes are both discernible and integrated. Discernibility means letting the player know what happens when they take action, and integration means tying those actions and outcomes into the larger context of the game.' (Ermi & Mäyrä, 2005).

This again displays similarities of demands that could equally be related to design theory and the technological affordance of devices for musical expression (Wanderley & Orio, 2002) (Tanaka, 2010). Where Wanderley & Orio's important considerations for HCI design regarding musical expression: 'Learnability' – i.e., the time taken to learn, 'Explorability' – the number of features and corresponding gestures or degree of gestures available to the user, 'Feature Controllability' – the resolution, accuracy, and range of perceivable features (Wanderley & Orio, 2002). Relate to Ermi & Mäyrä's 'discernibility'. Likewise, this is similar to mapping as described by Tanaka earlier in this chapter on page 33. Harvey compounds this further by stating, 'The stringency of rules in a game will dictate how game-like it is. Games that have flexible and fluid rules are more play-like… the unique space wherein the rules of the game in the moments of play supersede the rational rules of the real world.' (Harvey, 2009).

Flow

Though the study of games is a comparatively new research field, it does share some ground with longer-standing research fields. Subsequently, it is beneficial to compare related topics that are not exclusively specific to games. Csíkszentmihályi's theory of flow is ‘the psychology of optimal experience’, a concept which explores emotional and mental states such as ‘fun, happiness, satisfaction, and transcendence’ and how these states can be reached and utilised during a variety of activities ranging from sport and musical performance. (Csikszentmihalyi, 1990). Csikszentmihalyi defines total immersion and the corresponding pleasure as ‘flow’, which is measured between challenge and ability. Grimshaw, Calleja and Cairns dismiss ‘flow’ as being too vague (Grimshaw et al., 2011) (Calleja, 2011), and in the case of Cairns, it is ‘an almost holistic approach that defies proper analysis.’ (Cairns et al., 2014). They all agree that focused motivation in a task can be enabled through emotion.

Grimshaw accepts that Kearney’s work, which builds on Csíkszentmihályi's theory and suggests Sweetser & Wyeth’s work on ‘Game Flow’ does have merit. ‘Game Flow’ imposes a ranking based on a player’s response to eight different criteria (Concentration, Challenge, Player Skills, Control, Clear Goals, Feedback, Immersion, and Social Interaction) and is designed to have game-specific content. In the case of Kearney’s work on the studies and effect of ‘recursive loops of game-based learning’ (Kearney, 2007), this is unsurprising; ‘...the motivation for the player’s repeated engagement in the game-world is provided by immersion.’ (Grimshaw et al., 2011). This is not dissimilar to Cairn’s take that ‘for players to achieve immersion, they need to commit to playing the game, and it is only through taking the game seriously through emotional and cognitive involvements that they are able to achieve immersion.’ (Cairns et al., 2014). Additionally, Cairns refers further to his own work with Cox et al., in which participants were given two tasks, with one requiring more physical actions than the other. They found that ‘physical effort alone does not impact significantly on the level of immersion experienced.’ (Cox et al., 2012) (Cairns et al., 2014). This is further supported by Jennett’s work on measuring game immersion as a form of selective attention (Jennett, 2010). This is relevant to our research, as the same player demands may also be required for players to commit to their expressions. It also highlights that expression may not need to be overtly physical in order to impact the level of immersion.

Evaluating Player Experience

To determine the effectiveness of the gaming mechanics, we must use an appropriate research methodology for evaluating the player's experience. Denisova (2016) highlights the three most used questionnaires for qualitatively measuring player experience: the 'Immersive Experience Questionnaire' (IEQ) (Jennett et al., 2008), the 'Game Engagement Questionnaire' (GEngQ) (Brockmyer et al., 2009), and the 'Player Experience of Need Satisfaction' (PENS) (Rigby & Ryan, 2007), stating that these are often used as they are the most accessible for researchers and have all been externally validated. Specifically, the IEQ has been validated through producing repeatable results across a large study of two hundred and sixty participants, across diverse collection of different use cases, where factor analysis was conducted on the results (Jennett et al., 2008). Scale reliability was performed to ascertain the internal consistency of the questionnaire, with the IEQ yielding high levels of internal consistencies of 0.91 using Cronbach's alpha (Denisova et al., 2016).

The IEQ is comprised of five factors: cognitive involvement, real world dissociation, emotional involvement, challenge, and control. These are derived from the grounded theory study of Brown and Cairns (Brown and Cairns, 2004), 'which produced a strong fundamental measure in the gradation of immersion' (Nordin, A, 2014). 'The IEQ is a widely used questionnaire in determining the levels of immersion experienced by players. It has been tested much more empirically across a far-reaching array of different scenarios and game types' (Nordin et al., 2014), compared to other immersion and playing experience questionnaires such as the GEngQ and PENS questionnaire. This includes research into 'the role of challenge in producing a good experience' by Cox et al., where the effect of speed of interaction on immersion was tested. Along with, the effect of time pressure on immersion, and the interaction between time pressure and expertise on Immersion. (Cox et al., 2012). The 'Effect of Touch-Screen Size on Game Immersion' by Thompson et al., which investigated 'the influence on player's game immersion level by changing the size of the touch screen device used.' (Thompson et al., 2012). The study of 'Time Perception, Immersion and music in videogames' by Sanders & Cairns in 2010, which investigates 'if altering the degree of immersion in a videogame really does influence people's psychological perception of time passing.' (Sanders & Cairns, 2010). Alternatives include Sweetser & Wyeth's Game Flow questionnaire (Sweetser & Wyeth, 2005), Ermi and Mäyrä's questionnaire (Ermi & Mäyrä, 2005), and the 'Game Experience Questionnaire' (GEQ) (IJsselsteijn et al., 2013) which are not compared in Denisova's study as they were not readily available (although it is important to point out the GEQ is available and has been validated by IJselsteijn and Nacke). While all the individual questionnaires were designed to fulfil a specific requirement - for example, the GEnG 'was developed to assess the deep engagement of violent video game players' (Denisova et al., 2016), they all share considerable overlap with each other, as well as with the models of immersion as previously discussed. The biggest difference between the questionnaires is whether immersion is considered a component of the player experience, such as in the GEngQ or GEQ, or if it is considered the entire measurable experience, such as the IEQ and the one used by Ermi and Mäyrä.

As the questionnaire used by Ermi and Mäyrä is not readily available, we will first consider the IEQ. Jennett's IEQ comprises of five components: 'Cognitive Involvement, Emotional Involvement, Real Word Dissociation, Challenge, and Control' (Jennett et al., 2008) (Denisova et al., 2016). It is designed to be used across a variety of games (Cox et al., 2012) (Denisova & Cairns, 2015) and uses a five-point Likert scale to measure the player experience. It addresses components that are established in the cognitive basis and SCI model of immersion, such as imaginative and challenge immersion (Ermi & Mäyrä, 2005), as well as the psychological factors covered by Carr (Carr et al., 2006). Jennett states that 'through these experiments we have demonstrated that immersion can be measured subjectively (through questionnaires) as well as objectively (i.e., task completion time, eye movements).' The use of objective metrics to support the player experience questionnaire is a method we will discuss further below. However, Jennett does conclude that a common problem with evaluating through qualitative methods is difficulty in balancing the tasks in games of varying complexity (Jennett et al., 2008).

The GEngQ comprises of 'Absorption, Flow, Presence, and Immersion' measured in 19 positively worded questions on a seven-point Likert scale, with 'the questionnaire formulated in such a way that the engagement is a unidimensional experience, which ranges from immersion to flow' (Denisova et al., 2016). The component of 'Absorption', which we have not discussed outright yet, fits in with a cognitive basis of immersion and can be considered to address the issues of selected attention and player motivation specifically. However, as discussed previously, the notion of flow is problematic as it is far too vague as a concept and highly subjective. Consequently, data collected only from a player experience questionnaire of this type is not the most appropriate for allowing proper analysis. Similarly, the GEQ is used to categorise the player experience in seven different dimensions: Sensory and Imaginative Immersion, Tension, Competence, Flow, Negative Affect, Positive Affect, and Challenge. Unlike the cognitive basis and SCI model of immersion, which considers some of the affective experience as part of 'imaginative immersion,' the GEQ separates it into its own category. Although we are not primarily concerned with the affective experience, discounting corroborative data would be imprudent, providing it is meaningful data.

Metrics

As mentioned by Jennett (Jennett et al., 2008), immersion can be measured by combining the use of a questionnaire and objective data. The use of 'automatic event logging' is a long-standing practice (Hilbert & Redmiles, 2000), and in recent times has been used commonly for HCI behavioural observations, providing analysis of usability, affective state, and experience of the user (Nacke et al., 2008). These HCI methodologies have been adapted and applied to game testing (Pagulayan et al., 2012) (Pagulayan & Steury, 2004), as well as with combined attitudinal data (Nacke et al., 2009) (Kim et al. 2008). In studying player experience in games, this includes biometric/psychophysiological data (Nacke et al., 2011) (Nacke et al., 2008) and gameplay data (Nacke et al., 2009), which are logged as events within the game. Nacke combines biometric data from facial electromyography (EMG) and electrodermal activity (EDA), logging brain activity via an electroencephalogram to map player emotions, along with using eye trackers to assess fixations and assess cognitive processes. This is combined with a multitude of gameplay metrics to form:

> 'a valuable objective data source to user experience research and design, because these data offer quantitative, time-stamped information about the specific behaviour of players (DeRosa, 2007) (Kim et al., 2008) (Tychsen & Canossa, 2008). By combining

> metrical game data with other user experience measures… it is now possible to directly link game experience with game design elements' (Nacke et al., 2009).

The use of psychophysiological data by Nacke is an attempt to provide context and streamline the large number of gameplay events collected, addressing the issues highlighted by Kim (Kim et al., 2008) when applying classic HCI instrumentations to such large quantities of low-level event data.

Conclusion

As discussed above, within the topics concerning game design, literature related to novel use and approaches to game-specific design is fairly sparse. Literature for level design is either focused on aesthetics in a broad sense or methods of automatic content generation. Although there are a few studies with regard to the geometry of level design in relation to the structure of challenges, along with the importance of challenge for either immersion or for an engaging playing experience, there appears to be very little on the relationship between challenge and interaction structure. Similarly, sound design has literature covering methodology for situating the player within the game environment, but less so on techniques specific to games, as well as relating to methods of expression. Subsequently, the study will contribute to the field by expanding knowledge on the use of sound within game design. The aim is to demonstrate the novel use of sound in games, which highlights new methods or approaches for game design. By making prototype games and analysing and evaluating the playing experience, we can provide evidence as to whether expression with sound in games can enhance the player's experience. Existing literature reflects positively on this method of research and shows us there is potential value in what could be achieved as the domain of game design research is rather unpopulated, notably in the sub-topic of sound design. The research will also provide evaluation data, which could be useful in other research fields such as HCI, reinforcement learning, or gamification.

Research Methodology

Research Overview

This research is practice-based and explores novel sound design approaches and game mechanics 'so that areas of game design research can move closer to reflect game design practice' (Hook et al., 2017). This takes place over the series of three studies, each exploring a different gaming paradigm and addressing the research questions outlined in the introduction on page 20. Each study informs the next regarding both game and study design, following an iterative design process. Within each individual study, two similar versions of the game are developed and then tested in an A/B within-group format, which reflects typical industry practice (Siroker Koomen, 2015). The surveys used are refined between each of the studies and are comprised of a section devised by this researcher and a section that is the IEQ, a validated questionnaire for immersion (Jennett et al., 2008). This approach shapes our understanding of the playing experience and the player's sense of immersion in each game design. The initial preliminary study was performed in person, and the following two studies included remote and in-person participants. This change was necessitated by the COVID-19 pandemic and the desire to widen the number of suitable participants for the studies, increasing the diversity of the participants.

While each study takes place using a different game paradigm (a puzzle game, 3D platformer, rhythm game), the underlying gaming mechanic utilises the player's sonic expression to navigate the in-game challenge, and this remains the same in each game. Each of these different types of games typically has different factors to consider regarding game design, which will be discussed further in this chapter. All three games are heavily focused on musical rhythm and percussive event timing, as rhythm is fundamental to all musical (or sonic) events (Thaut et al., 2014). The first and second prototypes require the player to be proactive in their expression to be successful in the game. In the third game, the player's expression is reactive to the in-game challenge. Player interaction in the first and second prototypes also procedurally generates additional musical events. Player interaction also generates other game-integral environment elements required to navigate the challenge, along with game elements that support the challenge. In the second prototype, player interaction and movement are used to trigger Digital signal processing (DSP) techniques which impact the sound in the game. The third game focuses on synchronisation and the use of diegetic and non-diegetic sounds triggered by the player.

This practice-based research methodology made use of my decades of experience as a musician and a player to inform the game design. While each successive game was informed by the results of the previous game and resultant findings, they also drew on my experience as a practitioner of both disciplines. This can be considered as autobiographical design; as succinctly defined by Neustaeder & Sengers, 'Autobiographical design occurs when people build a system, use it themselves, use this experience to learn about the design space, and evaluate and iterate the design based on their own experiences.' (Neustaeder & Sengers, 2012). Throughout the development of all the games I frequently experimented within each game component (environment, controls, challenge, art, sound) to determine the most appropriate implementation for the game. Drawing upon two different practice backgrounds to inform my research methodology did cause some tension. For example, in the first prototype 'Boxel', different methods of interaction were tested for triggering the sound in-game. As a musician it felt important to have immediate sonic feedback when triggering a note, however as a gamer it felt punitively difficult if this was not sufficiently quantised. The solution was to iterate this method until both issues were addressed, triggering a sound at the point of interaction while simultaneously quantising the input, ready for the next time the sound looped around to that note. For each game in the early stages of development, this would start with sketches on paper, followed by individual developed components within a sandbox testing environment. For example, In the second game 'Red is Dead', a variety of different game environments were developed. All these environments focused on generated platforms based on musical input of a step sequencer but had vastly different behaviour; one where the platforms would recursively generate around the player's position, another where the platforms were fixed at the point of generation. As this game was a platformer it was imperative that it was clear how the player could interact with the environment and reach to the end goal. Subsequently, testing, evaluating, and reflecting on the playing experience was invaluable here. In this example there was a tension between being a developer and a player. As a developer, designing a game environment where platforms shifted and appeared recursively around the player felt novel and exciting, albeit with a more time-consuming development process than the alternate. For the playing experience it had a steeper learning curve and did not feel particularly intuitive. Whereas, for the version where platforms were generated in a static position it was less exciting as a mechanic/environment but was substantially easier to intuit as a player. After self-reflection and evaluating both approaches (and further experimentation of trying to improve and address shortcomings of each), the decision was made in favour of the more straightforward environment. As I needed the game environment to facilitate easily understood platforming for the player. A result entirely borne out of self-reflection and the autobiographical design approach, as described by Neustaeder & Sengers '…autobiographical design should rest on designers' genuine need for the system, which leads to a real engagement with the system that otherwise would be unlikely to happen.' (Neustaeder & Sengers, 2012). Autobiographical design has key advantages in that it builds on a highly engaged and expert view of pain points, that I have developed over years as a practitioner, highlighting areas of opportunity for development in a system. However, as an engaged autobiographical researcher it was important to continually reflect and be aware of any of my own biases and their own limits of insight into arising issues. I directly designed my research strategy to build on this, aiming to recruit participants with a wide range of different demographics and levels of experience in both music and gaming to ensure the widest possible range of engagement and input into my testing and iteration journey.

Technical Delivery

The two different A/B versions of the first study game were programmed in C# using Unity (Unity Technologies, 2004 - present) and Visual Studio (Microsoft, 1997 – present). Unity was originally chosen as it is easy to build for a variety of platforms, is well-supported with useful external libraries (such as assets/packages for a robust camera and input controller for Oculus Rift) and has well-populated Q&A help forums. The functionality to easily build to multiple platforms was necessary as the 'A' version was built for a mobile screen, whereas the 'B' version was built for VR using an Oculus Rift DK2 (Oculus VR, 2014 - 21). The A/B versions for both the second and third study games were programmed in JavaScript, making use of the A-Frame web framework (Marcos et al., 2015 - present) for 3D graphics and the

MaxiInstruments (Mimic Project, 2019 - present) library for audio. The A/B versions for both games were played in a Quest 2 (Reality Labs, 2020 - present) VR headset. The change from C# in Unity to JavaScript with A-frame was in response to the uncertainty in running in-person studies during the COVID-19 pandemic. While Unity can build to a web platform, the builds require considerable amounts of optimisation and have frequent issues critical to performance (CodeStarrk, 2021). A-Frame was the most viable alternative as it works within most browsers (including the commercial and affordable Quest 2 browser) and can be deployed for free and with ease within an HTML document. A-Frame provides 'a powerful entity framework that provides a declarative, extensible, and composable structure to Three.JS' (Cabello, 2010 - present), which affords comparatively high in-browser performance compared to Unity web builds. To generate audio and live synthesis (for the second and third games), MaxiInstruments was the most appropriate library to use, as the researcher had prior expertise in using the library (Maximilian (Grierson, 2009 - present)), albeit in the different programming language of C++ as opposed to using JavaScript.

To allow for rapid development and testing in the VR headset, the Accelerate Immersive framework Web App was used (Accelerate, 2022 - present). As this research focuses on novel game design and the resultant player experience, the practical component is game design as opposed to the programmatical approach. Subsequently, this will not be discussed here. However, each study discusses some of the techniques utilised for novel design and additional information can also be found within the support game design documents found in the appendix on pages 223, 237, and 250. As the second and third studies were remote, a small website was created to include all necessary details. This included participation consent, consent of data use / GDPR, game information, the game, a link to the corresponding Google Forms survey, and a final link to the remaining version of the game and survey to complete. This was hosted on Github Pages. Screenshots of the website for Study Two can be found in the Appendix on pages 247, 248, and 249. Screenshots of the website for Study Three can be in the Appendix on pages 260, 261, and 262.

Survey Design

The survey design for each study was developed iteratively in parallel with the game design. Over the course of the research, the survey design was streamlined to encourage participant engagement, reducing the overall time to complete while maintaining the same quality of results. For each study, the surveys used were delivered via Google Forms. In the first study, this was performed within the session in person. For the following two studies, this was done via a link on the study website. During the Preliminary Study, a pre-survey was used to ascertain the suitability of the participants. The pre-survey covered demographic characteristics and the gaming and musical experience of the player. It also had a pertinent exit question - the participants were required to have an interest in playing games to continue with the study. The participants were then given a short instructional video to watch and a set of four tasks to attempt to familiarise themselves with the game. They then played one version of the game (picked at random) before completing a post-survey. The post-survey was comprised of two individual sections. The first section was comprised of questions relating to playing experience, all written by this researcher. These were comprised of binary questions/answers, multiple choice, and five-point Likert scale questions. Such as 'Which sensory feedback was the most useful? Sound/by ear or Visual cues/by sight, 'How intuitive was the interaction?', and 'How clearly did you understand the effect your note placement had on your score?' The second section covered immersion and was comprised of the original 51 questions used to form the IEQ (Carr et al., 2006); this contained the final 31 questions selected for the IEQ (Jennett et al., 2008). A five-point Likert scale was specifically chosen over a seven-point Likert scale to increase the response rate and reduce the number of participants who dropped out of the survey (Buttle, 1996) (Babakus & Boller, 1992), as per the scale in the final survey (Jennett et al., 2008) It was hoped the additional 20 questions would elucidate more information regarding the playing experience in relation to immersion. Once the participants had completed the post-survey, they were asked to play the remaining version of the game, following the same process of practice tasks, playing the game, and completing the post-survey. In addition to completing the survey, participants also had their playing session recorded via screen capture. This was done so in-game playing behaviour could be observed and utilised for future game design choices. Lastly, participants were given an opportunity to ask the researcher questions relating to the research and provide informal qualitative feedback.

Many participants used the opportunity to provide feedback, highlighting how long the survey took to complete. Participants also reported the specified learning tasks prior to play were hard to complete, noting that they would prefer 'free' play to gain familiarity with the game mechanics. Participants also reported some redundancy within the questions asked in the post-survey and found some of the questions confusing. The following studies addressed these issues by removing the pre-survey and placing the demographic questions within the post-survey. Additionally, potential groups of participants were targeted due to their specific interest in games or experience with music. This avoided having an exit condition, as all participants were appropriate to take part. The number of questions was also reduced in each section. In the second section relating to immersion, only 32 questions were used - these were the finalised 31 questions of the IEQ with a single baseline question of player immersion on a 10-point Likert scale (Jennett et al., 2008). The first section was improved by removing similar-sounding questions that were covered within the IEQ, as these could be referenced independently in relation to the playing experience. This resulted in fewer complaints from participants, informally reporting they found the survey questions easy to understand. The final study was improved by increasing the number of questions in the first part of the survey relating to playing experience, with a greater number using a five-point Likert scale. This allowed for greater comparison between the two game versions and sections of the survey independently. Lastly, unlike the first two studies, participants in the third study were compensated for their time with an Amazon voucher. This was hoped to make participation a more attractive prospect.

Data Analysis

The survey results were exported from Google Forms as CSV, and the results were anonymised and cleaned in Microsoft Excel before being imported into SPSS Statistics. For each section of each survey, the results were checked for internal reliability (Hinton et al., 2004) (Tavakol & Dennick, 2011), and Cronbach's alpha was calculated. Where necessary, internal consistency was improved. As ‘Cronbach’s alpha is the most widely used objective measure of reliability’ (Tavakol & Dennick, 2011). ‘Crocker and Algina (Crocker & Algina, 1986) defined Reliability as the “repeatability of measurements for a qualification carried out on the same individuals under the similar conditions.” …The measurement results are as reliable as to the extent to which the results include less random errors (Cronbach, L.J., 1951)’. (Koçak et al., 2014). Subsequently, it is prudent to limit the number of items (in this case questions) where necessary. If the ‘…test contains many items that measure the same behaviour, deletion of certain items from the test would be an efficient way for attempting to improve reliability.’ (Koçak et al., 2014). In assessing the convergence of player questionnaires for games, Denisova states ‘scale reliability was performed to ensure internal consistency for each questionnaire using Cronbach’s alpha.’ (Denisova et al., 2016).

The data was checked for normality to determine the most appropriate test for comparing means (Löfgren n.d.) (Field et al., 2013). Due to the small sample size, Shapiro-Wilk tests were used for checking normality, showing if the data retained the 0.05 level of significance. If the necessary assumptions were met, the Dependent-means/Paired-Sample T-test was used to compare means. The effect size was also calculated using Cohen’s d (Wiseheart n.d.) (Cohen, 1988). This is because Cohen’s d is designed to measures the size of the difference between two different groups. The effect size is independent of sample size and can be used to quantitatively compare results of studies performed in different settings. Where the small sample size could be affecting the normality of data, the non-parametric Wilcoxon Signed-Ranks Test was used to compare means (Field et al., 2013). In addition to analysing the whole survey combined per game and each individual section, the results from the IEQ were reported, broken down into the five immersion factors of cognitive involvement, emotional involvement, real-world dissociation, control, and challenge (Jennett et al., 2008). Demographic characteristics such as prior gaming experience and musical experience were also observed in relation to playing experience and immersion. Once all sections of the survey were analysed, the findings were evaluated. The outcome of the evaluation resulted in further questions pertaining to game design, some of which filtered into the following study. These were determined based on feasibility, scope, and relevance to the research questions. Where appropriate, the evaluation identifies other potential areas of further research that do not fall within the scope of this research.

Game Design Overview

As discussed in the research methodology earlier in this chapter, the in-game challenge will situate the embodied sound and game action together. This places specific requirements for how the game environment, controls, and in-game challenges are designed. Within each game, it must be clear how the player can interact within their environment, how the same interaction generates sound and the impact this interaction has on the player completing the in-game challenge. For each of the three prototypes, the demand on game design will be different relative to the specific gaming paradigm of that game genre. For example, in a rhythm game where the player shoots targets, a demand would be the sequencing of targets that exist in the game environment for the player to shoot that generates a sound upon collision. These targets would need to appear and disappear at highly specific intervals relative to the game music. This is different to the demand of a typical 3D platform game, where the player is often given as much time as needed to interact or navigate the game environment. However, in identifying these differences early in the design stage, it was possible to align the demands between each prototype as closely as possible. This was beneficial in mitigating the amount of time spent in development for each game. It also allows for clear iterative design, as knowledge regarding game design from the first prototype can be directly applied to the second and third prototypes. By focusing narrowly on very specific and limited methods of sonic interaction pertaining to rhythm, a clear singular general game design approach could be maintained across each prototype, despite each one existing in a different genre of game. This approach was taken as rhythm is central to any musical experience (Thaut et al., 2014). Further detail on how this was accomplished technically for each prototype can be found within their respective study chapters and within the corresponding game design document found in the appendix on pages 223, 237, and 250.

Game Design: Challenge

For each of the three games (puzzle, platformer, rhythm game), the in-game challenge was distilled down to the lowest abstraction possible for that specific genre of game, albeit incorporating a gaming mechanic that was conducive to the necessary interaction for sonic expression to occur.

Boxel – Puzzle Game in Preliminary Study

In the puzzle game, players were asked to score as many points as possible by making a groove out of four percussive samples. The level took place over a set duration relative to the musical meter employed within the game. The players were not explicitly informed how the game awarded points (measuring note distribution and self-similarity across metric layers to determine coherency); they were instructed that they 'score points by building a coherent rhythm' (page 228) and were left to experiment on their own. This encouraged experimentation and exploration of the gaming mechanic that assessed their arrangement over time. Players were not given the option to delete a bad input, and the game would remove points for 'bad' input or the lack of input (the minimum requirement of input was four bars of music). This was a deliberate design decision to encourage players to be thoughtful about their input while remaining under pressure to complete the task. Time-sensitive tasks feature prominently in puzzle-orientated games, such as Minit (Devolver Digital, 2018) or Candy Crush Saga (King, 2012), and pressure plays an important role in game design (Hills, 2022).

Red is Dead - 3D Platform in Study Two

In the 3D platformer, players were asked to navigate from a starting point to an endpoint while avoiding enemies. Like the puzzle game, this challenge required the player to generate a rhythmic arrangement from four percussive samples to generate the necessary platforms (and corresponding enemies) to reach the goal. Unlike the puzzle game, there was no time limit. However, as each enemy moved around the game environment in a loop, players could not stand still indefinitely on platforms. The only way to engineer space to avoid the enemies required the player to make sounds and interact with the game environment. This facilitated the necessary pressure to encourage players to continue their interaction. The players were also required to collect tokens to generate platforms. The platformer also punished the player by positioning the tokens in harder-to-reach places if the player's input was found to be incoherent. This technical approach for this was similar to the puzzle prototype, where players were awarded points for coherency. However, in this instance they were punished for a lack of coherency as opposed to rewarded for coherency.

Grail – Rhythm Game in Study Three

In the third prototype, players were asked to shoot targets with lasers before they hit the player. The targets were arranged in 3D space relative to the musical arrangement of the bassline for the game music. They moved in time with the backing music and were generated in positions based on their pitch. Upon a successful shot, the collision triggered the corresponding bass note and destroyed the target. Players were instructed to shoot all targets and reach the end of the level without being hit multiple times, as this would end the game prematurely. Unlike the first two prototypes, the player is not responsible for the overall arrangement of the rhythmic events. However, they are responsible for the timing of each note, where even a small change in timing (triggering the note early or late relative to the intended arrangement) leads to a noticeable impact on the musical output. This is different to the first two prototypes, where the player's mistakes pertaining to sonic expression were punished relative to the gaming challenge, as opposed to not making a significant impact on the outcome of the gaming challenge but impacting the overall sonic experience greatly.

Game Design: Controls

For each prototype, appropriate consideration regarding interaction and controller/device was made for each specific platform the game was delivered on. All three studies had at least one VR game. Subsequently, attempts were made to follow typical button layout conventions for the genre of the game.

Boxel – Puzzle Game in Preliminary Study

For the VR version of the puzzle game, it was appropriate to have the camera position controlled by the headset position and orientation, allowing the players to move around the puzzle, as the interactive graphics were delivered in a fixed position.

Red is Dead – 3D Platform Game in Study Two

The movement approach taken in the Preliminary Study was not appropriate for the platformer as it would require an omnidirectional treadmill such as a Virtuix Omni (Virtuix, 2017), an expensive device not available to many players. The solution was to use the thumb-sticks on the Oculus Touch Controllers for player position, while maintaining the orientation from the facing direction of the VR headset. Both 'X' and 'A' buttons were used for jumping, allowing the player to perform this input with either controller. Likewise, the input 'Y' or 'B' buttons were used to destroy a previous inputted platform.

Grail – Rhythm Game in Study Three

The rhythm game made use of the headsets facing direction as described for the 3D platformer to determine the camera orientation. The controllers' main triggers were specifically used as the input method closely resembles pulling a gun trigger (compared to a thumb press on a flat button that would be the alternative input available).

Game Design: Environment

The visual game environment for each prototype was pared down to its most abstract form. This was a deliberate decision to limit elements outside of those I wished to study, factoring in only those essential to the paradigm of the game. All three games took place in a 3D game environment. Each prototype consisted of a single level to play. There were no in-game instructions or tutorials, with these being delivered prior to play. Each game had a different game environment. However, there was continuity in some design elements regarding art and sound design.

Game Design: Art

For each game, the art design is deliberately clean and concise. It is deliberately minimal to not distract from the sound in the game, as sound is the salient feature within my research. Simple, bold colours and the use of transparency are used to differentiate between different game elements.

Boxel – Puzzle Game in Preliminary Study

In the puzzle game, having a clear colour palette helps identify the cube pieces the players interact and create. Each step of the rhythm sequence is a single cube displayed on a two-by-two cube grid on the x-z axis. These are placed inside a larger semi-transparent two-by-two-sized cube to highlight that each cube is itself a subdivision of a division of a beat and bar, with four of these larger divisions displayed on the screen at any one time. The incoming steps are displayed on the y-axis above the corresponding coordinate for the note. At the end of each completed level, the player's complete rhythm structure is highlighted within the four larger semi-transparent cubes before fading out as a new level starts. Additionally, the background has a two-tone colour gradient moving from the incoming level at the top of the screen to the current level background colour at the bottom; as the player moves through each repetition of the musical loop, the gradient moves lower until the screen is filled with the new colour, implying the player's position in the level.

Red is Dead – 3D Platformer in Study Two

In the platform game, simple, bold colours are used to differentiate between the different platform types and the collectable tokens, and it has a clear message that red objects are to be avoided. It was necessary to give the playheads some transparency so players could see where the tokens were through them.

Grail – Rhythm Game in Study Three

The rhythm game uses simple, bold colours to differentiate between the different game environment elements and the target blocks for the players to shoot. The corresponding lasers are clearly coloured so players understand they can only shoot certain targets with certain lasers. The colour palette was chosen by this researcher to complement the timbral qualities of the music presented.

Game Design: Sound

Electronically generated samples were chosen for each game. This was a deliberate choice by the researcher, as it lends itself naturally to Digital signal processing (DSP) techniques from player interaction. For the platformer and rhythm games, these samples neatly accompanied the player-controlled and parameterised synth used to generate sound 'live' in the browser.

Boxel – Puzzle Game in Preliminary Study

In the puzzle game, the four percussive sounds are sampled from a Roland 909 (Roland, 1983 - 1985). This gives the player the option of triggering a kick drum, a snare drum, a cowbell, or a hi-hat sample with corresponding button presses (in the VR version) or swipe direction (in the screen-based version). The backing track is arranged based on the player's previous rhythmic input. Every time the sequence loops the previously inputted player rhythm is used to arrange the backing music arrangement. This is achieved with a ruleset based on the distribution of notes and accents within the previously generated level. As the backing track is sample-based, it allows for quick changing of sounds, meaning each level could have its own unique sound in terms of style and timbre. However, these will be limited to sounds clearly electronic in origin – serving both as a choice of aesthetics and to embrace the origins of the gaming mechanism – the step sequencer, and to hide the limitations the sequencing mechanism currently has – such as not being able to account for micro-rhythms of steps, swing, triplets or other non-even note divisions.

Red is Dead – 3D Platformer in Study Two

For the platformer, the four percussive sounds are sampled from a Korg ER1 (Korg, 2003). Using four buttons on the controller, the player could trigger a kick drum, a snare drum, a heavily filtered high-pitched 'bleep' hi-hat, and a low-pass filtered 'bloop' cowbell. These were deliberately chosen for their suitability for the desired retro-inspired glitch / IDM style. This style was chosen as it lends itself to the rapid changes of parameterised effects that are applied from the player's movement. As mentioned briefly above (and in the game design document in the appendix on page 237). player movement and interaction change the parameters of the backing synthesisers. These synthesisers are triggered 'live' in the browser using the MaxiInstruments library (Mimic Project, 2019 - present). Player movement, such as jumping, changes the sample rate playback, which adds interesting timbral changes when applied ad-hoc to our percussive samples. This is a typical feature of our chosen style. Additionally, the platform game has a bass synth with parameterised Low Frequency Oscillator (LFO), filter frequency, and pitch modulation. There is also a midrange backing synth with parameterised features and a synth for the collision of our enemy playheads and players, which triggers a note from an array. These are subject to audio effects such as delay and reverb, which are also parameterised based on in-game activity. In the 'B' version of the game, the rhythmic sequence (and subsequently platform structure) has been designed to create a resultant rhythm that is coherent with the style of the chosen electronic music genre.

Grail – Rhythm Game in Study Three

In this game the player triggers either a musical note or a sound effect. For the 'A' version, when the player successfully hits the target the bass synthesiser is being triggered in the browser via MaxiInstruments. This was beneficial for triggering notes with accurate timing (given this is a rhythm game, timing accuracy is crucial to the playing experience). The remaining backing music exists as a sample being triggered once upon start. The sound effect for when the player takes damage is a short one-shot sample, originally synthesised by this researcher using an ES2 in Logic Pro (Apple, 2002). In the 'B' version, when the player successfully hits the target it triggers a sound effect, which is another one-shot sample, originally synthesised by this researcher using the same ES2 plugin in Logic Pro. The backing track contains the original bassline. The sound design for the rhythm-game was a deliberate choice of techno as a genre/style due to the (typically) relatively slow bpm compared to other genres of dance music. The backing track has a bpm of 127. Stylistically, the music strafes minimal and progressive tech, in homage to such artists and tracks as '0800' by Extrawelt (Schaffhausen & Raabe, 2006), 'Some Polyphony' by Petter (Nordkvist, 2006), 'With You' by Ricardo Tobar (Tobar, 2008), 'Isst' by Tiefschwarz (Schmalbach, 2005) and remixed by Nathan Fake (Fake, 2005), and 'The Was Pink' by Nathan Fake (Fake, 2004) and remixed by James Holden (Holden, 2004), along with other similar works of dance music found on the record label 'Border Communities' (Holden, 2003 - present). This style also lent itself to changing the distribution of notes, progressively getting more complex as the music developed over the course of the full track. Additionally, this also allowed this researcher to keep the instrumentation down to three synthesisers and a drum machine.

Preliminary Study: Comparing the Playing Experience of VR and Mobile Platforms when using Sonic Expression as a Gaming Mechanic

Abstract

This chapter presents a preliminary study of player experience when attempting to use musical expression as the primary gaming mechanic. It explores the player's experience and sense of immersion across two different gaming platforms. This will help determine if the platform and its impact upon the experience of playing the game where musical and sonic expression is a core game mechanic. The takes place in a prototype puzzle game called 'Boxel', where players develop a coherent rhythmic arrangement to earn points. It is hoped that the outcome of the study will help establish whether there is any advantage in developing music games specifically for virtual reality (VR) over traditional screen-based platforms. It will also help to establish what impact the use of sonic expression can have on immersion and the overall playing experience, specifically in relation to the challenge component.

Introduction

We conducted this study to understand more fully if either sonic or musical expression was more effective as a game mechanic in either standard 2D screen-based games or, alternatively, within VR. The study was a 'within group' study. There were 12 participants. The study focused on the exploration of the difference in the player's experience and the player's sense of immersion between screen-based and VR in the specific context of a music game, where the goal of the game was a form of musical interaction that can be thought of as sonically or musically expressive. The rationale behind the study is to inform game design. If the playing experience greatly differs between the two platforms, then further works as part of this research will be developed exclusively for the most appropriate platform. The outcome of this preliminary study helps to determine if VR is a platform better suited for a specific type of musical interaction game than traditional screen-based platforms. The study highlights important conditions for interaction in an expressive musical experience, an enjoyable player experience, and an immersive experience. It was decided at the outset that if there was no discernible difference, then developing games for screen-based platforms would instead be the right course of action.

This has important implications for development with respect to musical and/or sonic expression. The technical advantage of having a small portable device that is networked opens the potential for different design choices, such as local multiplayer with a high number of players. This would be more difficult to achieve with VR due to prohibitive cost, space, and computational demand. The technical qualities of a mobile device are also more flexible in terms of input methods offered. Most mobile devices have a variety of sensors, such as accelerometers, that could be used for player input, affording novel game interactions. They also have touchscreens which offer a range of different user interface options, such as sliders and dials; these can be tailored for the game. When designing for VR, you are limited by what the physical controller allows users to achieve. These become interesting potential options that could lead to mobile platforms being a better option for a game focused on musical expression, as opposed to VR.

Further to this, if VR was proven to be the more suitable platform for sonic/musical expression as a game mechanic, then further development would only be in VR. This opens further topics of conversation regarding the quality of embodied interaction that takes place, the player's perception of the task, and the relationship with expression. Both potential paths have the capacity to yield interesting results regarding game design, the relationship between music performance and game challenge, and the effects this has on the immersive playing experience.

Methodology

The first step was to make a game that contained appropriate features to test musical or sonic expression as a gaming mechanic. The first feature is a simple interactive method for the player to generate sound. The second feature is a challenge for them to complete, with a corresponding reward. The first is necessary as without any method for the player to generate sound, we would not be able to answer our question. The second is necessary as without a challenge for the player to overcome, it would not be a game but a toy. The game is a simple puzzle game and was built in Unity using C#. The in-game challenge is designed in such a way that it can be solved through a multitude of different inputs. A successful order of sequence is entirely dependent on the player's inputs. The study A/B tests the game on two different gaming platforms within-group. The two versions of the game were similar, with appropriate changes made to cater for the different interfaces and necessary in-game camera components. A pre-survey was run to determine if the participant was suitable for the study. The study was run in person, with a post-survey completed after each play-through. Each participant played both games in a random order to limit bias, as players potentially could gain experience between sessions.

Game Design and Dynamics

The game introduces the player to broad concepts relating to musical rhythm. The player is instructed to build a rhythm and the challenge of the game is to score as many points as you can. In the screen-based version, this is by swiping in a direction, with each direction producing a different percussive sound (swiping down generates a kick drum sample, swiping up generates a snare drum sample, swiping left generates a hi-hat sample, swiping right generates a cowbell sample). In the VR version, this is accomplished by placing your hands in a specific 3D space and pressing one of the four trigger buttons (each with a corresponding sample – the same four samples as described above for the screen-based version). Each time the player triggers a sound, it gets added as a note in that position within a four-bar loop. The players are given four repetitions of this four-bar loop to add to their rhythm. Each successive bar input is checked against the previous bar, comparing the distribution of notes per subdivision per bar as it moves through the loop. At the end of the four-bar loop, the salient features of the loop are used to generate the backing track for the proceeding four-bar rhythm, and the user starts again on a new level as their previous input is deleted over the period of eight bars. This is considered a game 'level'.

The process for checking the user's input and their generated rhythm is as follows:

1) The player adds a note to the sequence
2) The position of this note is checked relative to the sequence of notes for the backing track
3) The position of the note is checked against the previous beat
4) The position of the note is checked against the previous bar of see if the input is a rotation or inversion of the previous rhythm
5) The distributions of notes for the current bar and previous bar are checked
6) The symmetry of notes within two bar groupings are checked
7) The symmetry of notes is checked across the full four-bar loop
8) The player is awarded points or has points removed from their score based on the outcome of the checks above

This process occurs for every input by the player. Figures 1, 2, 3, and 4 show a player triggering notes in the VR version of the game.

Figure 1: Boxel VR: The arrangement is empty, and the player's hand is outside the trigger area.

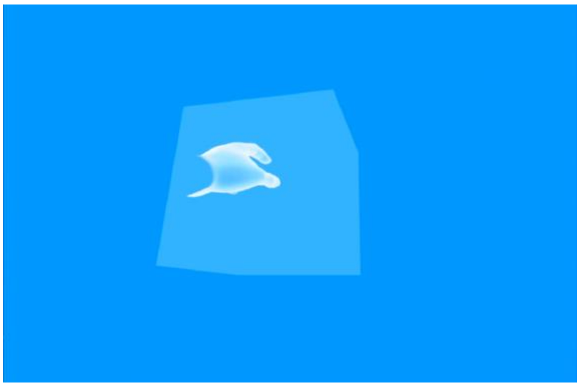


Figure 2: Boxel VR: The player's hand is now in the trigger area and can place a note with a button press.

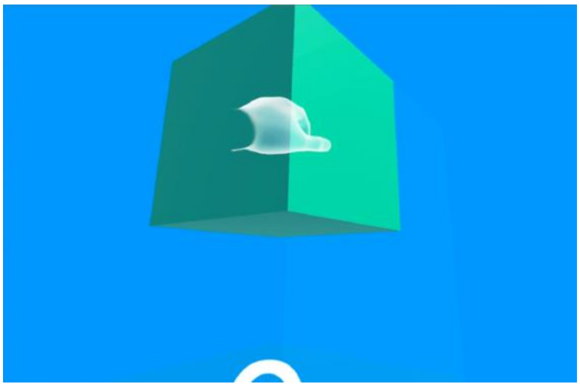


Figure 3: Boxel VR: The player has triggered a note.

Figure 4: Boxel VR: The player has triggered many notes!

To make progress within the game, the player must continually add notes to the loop. The notes added must relate in some way to the previous notes added. How coherently the notes relate to one another determines how many points the player scores. For simplicity's sake, we are defining coherence to be pattern matching of small rhythm fragments from the previous rhythm to the newly created one, the similarity of note distribution between subdivisions relative to the beat, the bar, and the four-bar loop, and how these note distributions imply or characterise the metric pulse (either ‘pushing’, ‘pulling’, or explicitly stating the emphasis of the beat). The coherency of arrangement was measured across relative metric subdivisions for each of the four possible percussive elements, along with measuring the coherency of the overall rhythmic arrangement. This was accomplished by analysing a weighted distribution of the player’s percussive elements, along with analysing the self-similarity across each metric subdivision. For the player’s input to be considered coherent, it must deviate sufficiently within at least one of the metric layers. If there is not sufficient input per bar or per repetition, the player’s rewards diminish, ensuring that players are rewarded for novel interactions. If the player’s input stagnates over two of the four-bar loops, the player loses the game. Additionally, if the player’s input is not coherent across the four-bar loop twice in a row, the player also loses the game. In essence, the game is about how a player chooses to develop a musical rhythm over time, rewarding a methodically measured response.

The game environment shares similarities with a traditional step-sequence synthesiser. However, rather than displaying each step of the sequence at once with a moving ‘playhead’, the game environment uses a fixed camera position acting as a playhead and instead moves the steps based on their position relative to the playhead accordingly. This allows the player to make cursory glances at incoming structures whilst giving them a consistent visual location of where the playhead is (at the top of the screen). The sequencer also exists in 3D, with the different sounds arranged across the x- and z- axes evenly. As the player activates a step in the sequencer, they are simultaneously building simple 3D structures comprised of these cubes. This provides visual feedback and reinforces the player’s understanding of their generated rhythmic structures.

Both versions of the game were identical, albeit with some differences inherent to the respective platforms (see figures below). For example, the camera positions are different. In the screen-based version, the camera is in a fixed position. In the VR version, it is attached to the player's head position and can move in any direction.

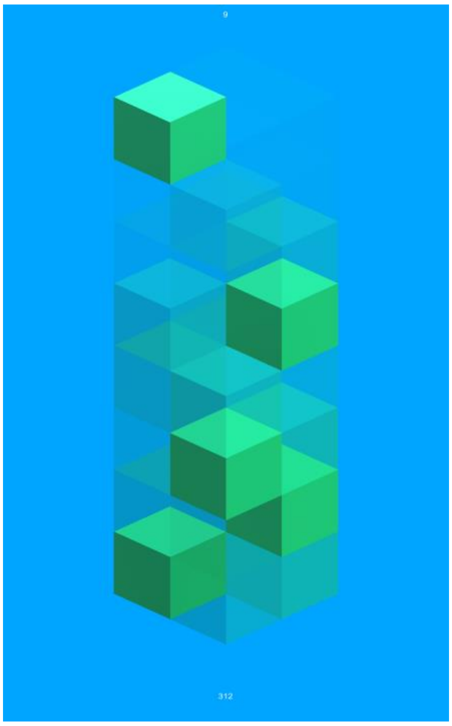


Figure 5: Boxel Mobile: Camera fixed, swipes create cubes at the top of the tower.

Figure 6: Boxel VR: Camera moves based on head position of the player.

Figure 7: Boxel VR: player triggers sound by placing hands in top quadrant.

As the VR version allows the player to move, there is also a graphical representation of where the player's hands are in the game environment. This is not necessary in the screen-based version, as the player never moves. Due to this movement, there is also a different input method. The screen-based version requires the player to swipe with their finger in a direction along either the x or y-axis of the screen to generate input with different sounds being produced for up, down, left, and right. At the same time, the VR version requires the touch controller button to be pressed in a specific 3D space within the game environment to generate the sounds. Both then add these sounds to a looping sequencer in the temporal position they were added. To visually represent the player's hand in the activation space for the interaction, there is a visual cue of a semi-transparent box around their hands. This is not found in the screen-based version, as the player is always free to make their input. Similarly, as the player is required to move around the game environment in the VR version, a floor was added, along with a layer of low-moving smoke, to help differentiate their y-axis. This is not found in the screen-based version, as it was not needed. Once the game was built, each version of the game was provisionally tested. Established player experience questionnaires that are already validated were reviewed, and appropriate questions were decided. Many of the questions came from the immersion questionnaire (IEQ) (Jennett et al., 2008). This is because the IEQ is the most relevant verified questionnaire, containing questions specifically relating to games. Participants then played both games, completing the same survey after each one.

Potential participants were asked to complete a pre-survey to determine their suitability for the study. To be eligible for the study, participants must either have previously played video games or have a desire to play video games. If they had previous playing experience, they were asked questions regarding how frequently, on what platforms, and whether they had any experience with any existing music or rhythm games. They were also asked questions relating to their prior musical learning. The pre-survey was created using Google Forms and was sent out to candidates via email. For an example copy of the pre-survey questions, see the attached document Pre-Survey Example (p 216).

After signing a consent form, the participants were asked to play two different versions of the same game, one screen-based on an iPhone, the other a virtual-reality game running on a PC using an Oculus Rift with touch controls. Before playing for the first time, each player was shown a short introductory video that outlined the main game objective, the interface, and the game environment (see the attached document for the introduction script of the video, p 215). This allowed each player to go into the training and challenge tasks of study with the same level of knowledge as each other.

The training task consisted of five sequential steps to build a simple drumbeat, which would score in-game points, and to give the player a 'hands-on' introduction to the gaming interface. Each participant was allowed no more than five attempts at the training task, although they were also given a couple of minutes to freely experiment with their input to feel more comfortable with the interface if they required it. Participants were then asked to play the game, performing the following three tasks:

- They were encouraged to score as many points as possible within four levels.
- They were encouraged to score as many points as possible with four levels using one sound.
- They were encouraged to score as badly as possible over three levels - deliberately trying to lose the game after one level of trying to score points.

The challenge tasks were randomised per participant per group to avoid any potential bias. Upon completion of the three challenge tasks, they were then asked to fill out a post-survey, evaluating their playing experience. Each group would use the same post-survey.

As players might find the game less challenging or more intuitive once they have gained some playing experience, regardless of whether the version of the game is different, the order in which participants played the different versions of the game was randomised to limit any potential bias. For the VR group, players were then introduced to the headset, the controllers, and the 'Guardian Player Protection System' and had the equipment calibrated for their height and head size. The 'Guardian Player Protection System' is the software designed by Oculus to warn the player if they reach the edge of their play area. Before playing any game, the player sets a boundary that the game understands as the edge of the play area. Should a player move outside of this space whilst playing, the controller vibrates, and visual crosshatching appears in the headset to warn the player.

The post-survey was comprised of three sections. The first contained 16 questions, 15 of which were on a five-point Likert scale. One was a binary choice. The questions were written by this researcher. These covered the broad playing experience and design of the game, specifically the player's interaction. The second section was a further 50 questions on a five-point Likert scale that contained Jennet's 31 questions from the IEQ (Jennett et al., 2008) and an additional 19 questions used in testing for the IEQ by Cairns. The IEQ is independently verified, making it a suitable choice for post-survey questions regarding how immersed players felt. It comprises of five sub-scale components: Cognitive Involvement, Emotional

Involvement, Real World Dissociation, Challenge, and Control. The third and final section was an optional comment section in which participants could offer any further insight regarding their playing experience. Screencasts of the player's performance were also recorded from the VR unit, although these were not for the purpose of formal analysis. The surveys were completed via Google Forms. For an example copy of the questions, see the 'Post-Survey Example' attached (p 220). Players were rewarded for completion of the survey with chocolates and a short break before switching over to the other group and continuing the study.

Each group had 13 participants. Unfortunately, due to intermittent Wi-Fi, one participant's post-survey results did not save correctly (discovered at a later date). A 14th person was only able to complete one of the group studies, but due to technical issues, they were not able to complete the second group study. Subsequently, these two results would have to be removed as a within-group study requires both groups to be successfully completed by each participant. This leaves a total of 12 participants. The removal was done in Excel once the data was exported from Google Forms in CSV format.

Both group post-surveys were combined with an additional group label variable added to identify which group the results related to. The binary design question 'Which sensory feedback was the most useful?' was removed, leaving only questions with a five-point Likert scale. The post-survey was then split into three different versions. One comprised only the survey questions, one comprised only the IEQ, and the third with both combined. It is also worth noting that for questions 6, 8, 9, 10, 18, and 20 of the IEQ, a negative scale was used, so the results for these questions were reversed before being exported. After the data was cleaned, it was imported into SPSS.

For each version of the survey data, Cronbach's alpha was calculated to measure the internal reliability, as described in the Methodology Chapter on page 60. Cronbach's alpha for all 65 questions was 0.939, showing excellent internal consistency. Part one of the survey had a Cronbach's alpha of 0.774, showing acceptable internal consistency; part two had a value of 0.929, showing excellent consistency and the 31 IEQ questions within had a value of 0.872 which is good. By comparing Cronbach's alpha against each individual question's value, the following questions could be removed as they showed the alpha value would increase if deleted (across each measure taken of the Cronbach alpha). This improved the reliability of the questionnaire as it removed items that were found to measure the same behaviour (Koçak et al., 2014).

Q3: How Challenging was the game?
Q8: How aware were you of the changes in the accompanying background?
Q11: Regarding the visual feedback and your position in the music loop, how important was the bar counter?
Q12: Regarding the visual feedback and your position in the music loop, how important were the moving semi-transparent cubes?
Q31: To what extent did you feel as though you were moving through the game according to your own will?
Q32: To what extent did you find the game challenging?
Q35: To what extent did you find the game easy?
Q48: Things seem to happen automatically
Q50: I feel scared
Q53: I get wound up

After removing these questions, the Cronbach's alpha for the remaining 55 questions was 0.954. Whilst this is excellent internal consistency, it is worth considering that the alpha could be high as a result of redundant questions (Tavakol & Dennick, 2011).

The results were then totalled and divided by the number of questions (55). Totals for each separate section of the survey were also generated. Section one (consisting of my questions) is the first 11 questions, and section two is the remaining 44 questions. A separate total for the finalised IEQ was also provided, which consisted of the first 28 (three questions were removed after calculating Cronbach's alpha) of section two.

Data Analysis: Combined Data

To determine the most appropriate test to use for comparing means, the distribution of the data must be checked for normality (Löfgren n.d.) (Field et al., 2013). This is done by first approximating the distribution by checking the histogram and Q Plot for the two separate groups by eye. The two histograms and Q Plots can be seen in the figures below.

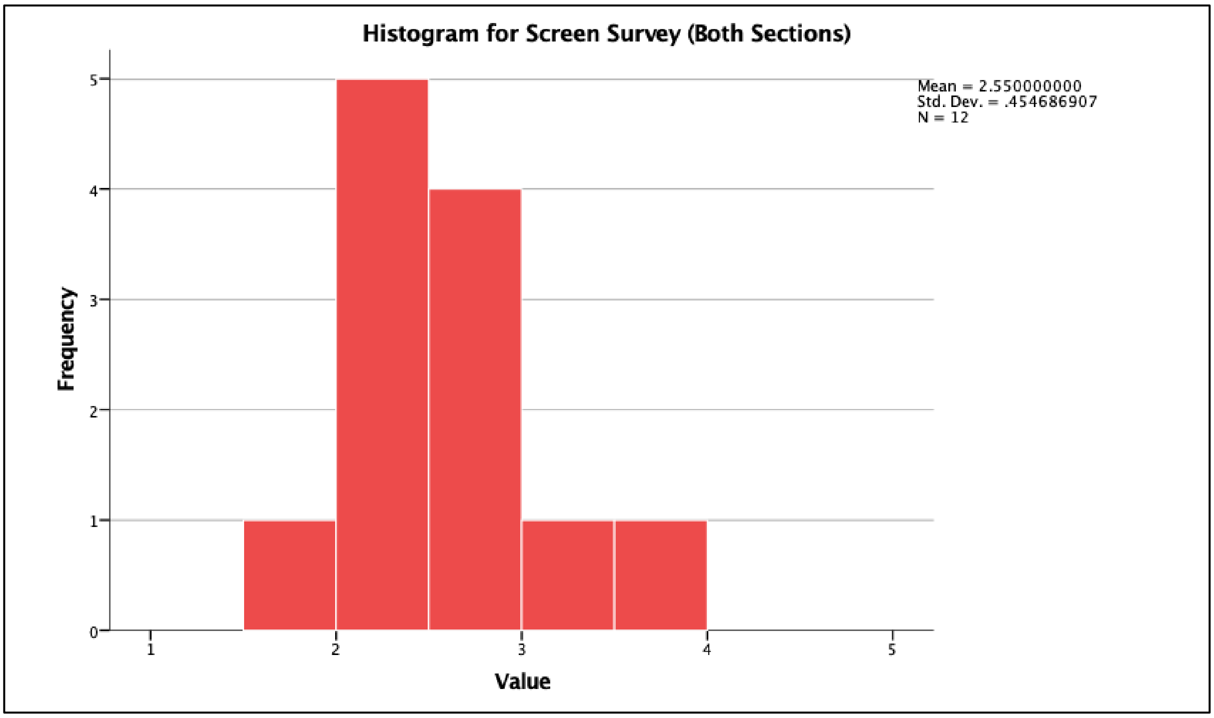


Figure 8: Preliminary Study Game A (Screen Version) Histogram Results for Complete Survey

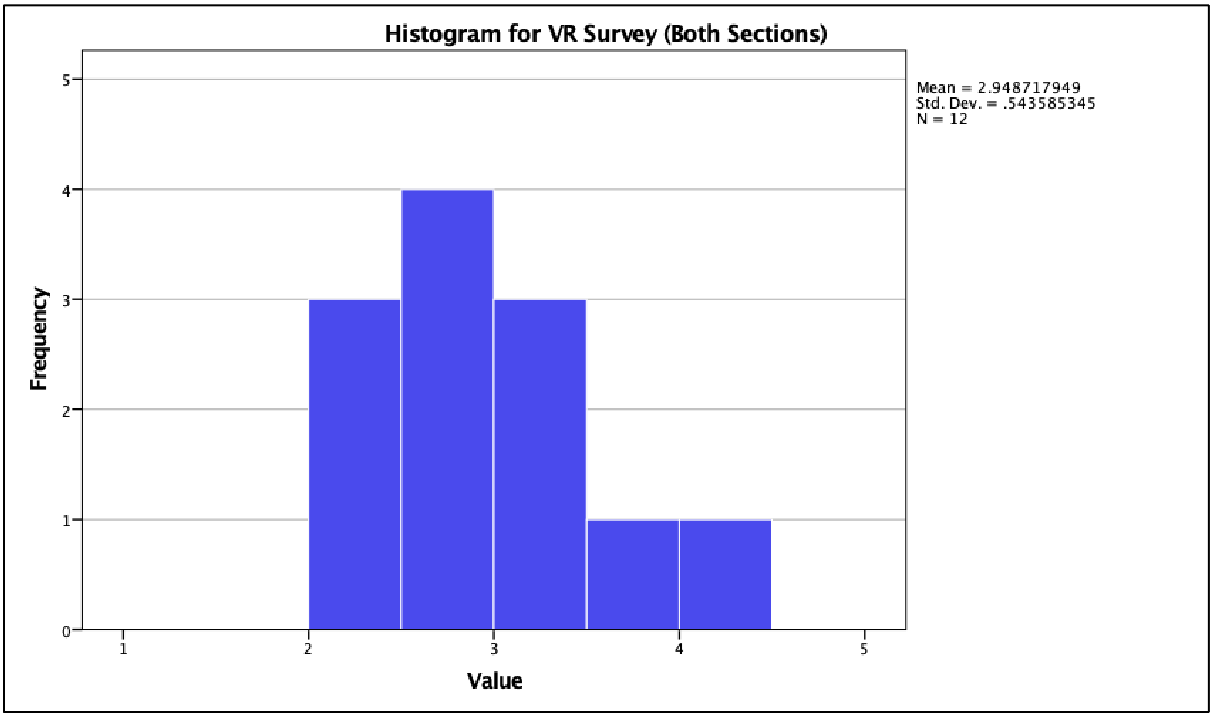


Figure 9: *Preliminary Study Game B (VR Version) Histogram Results for Complete Survey*

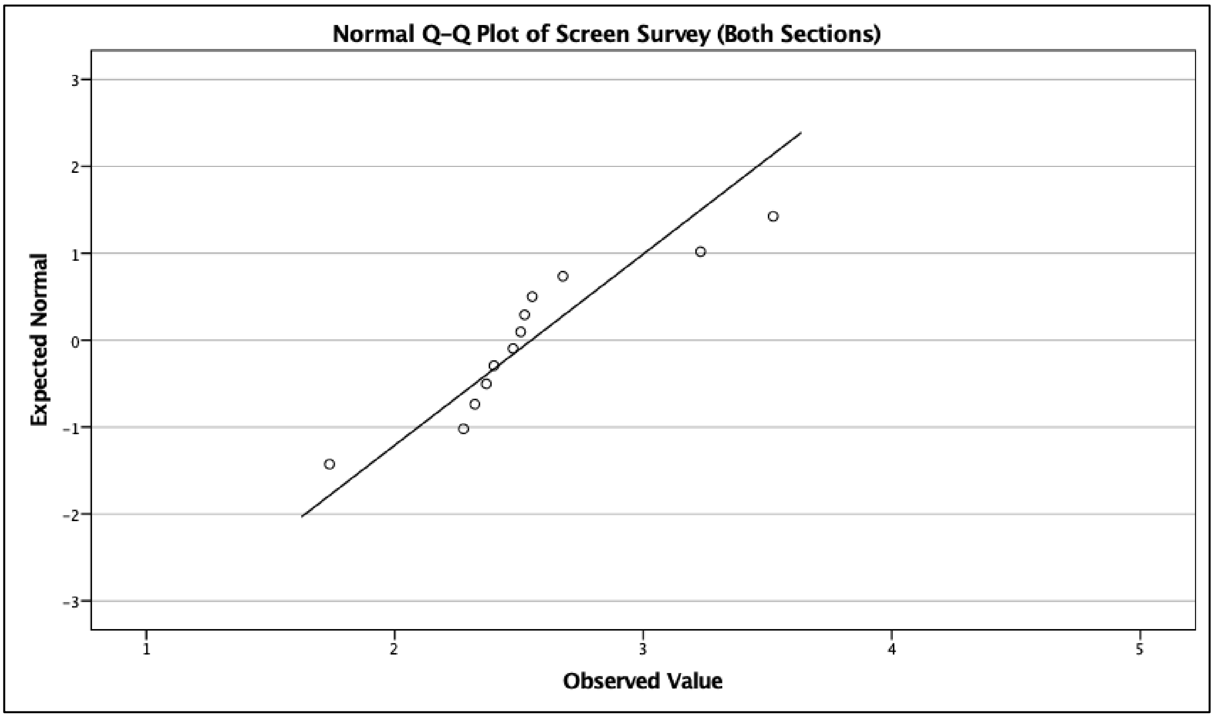


Figure 10: Preliminary Study Game A (Screen Version) Q Plot for Complete Survey

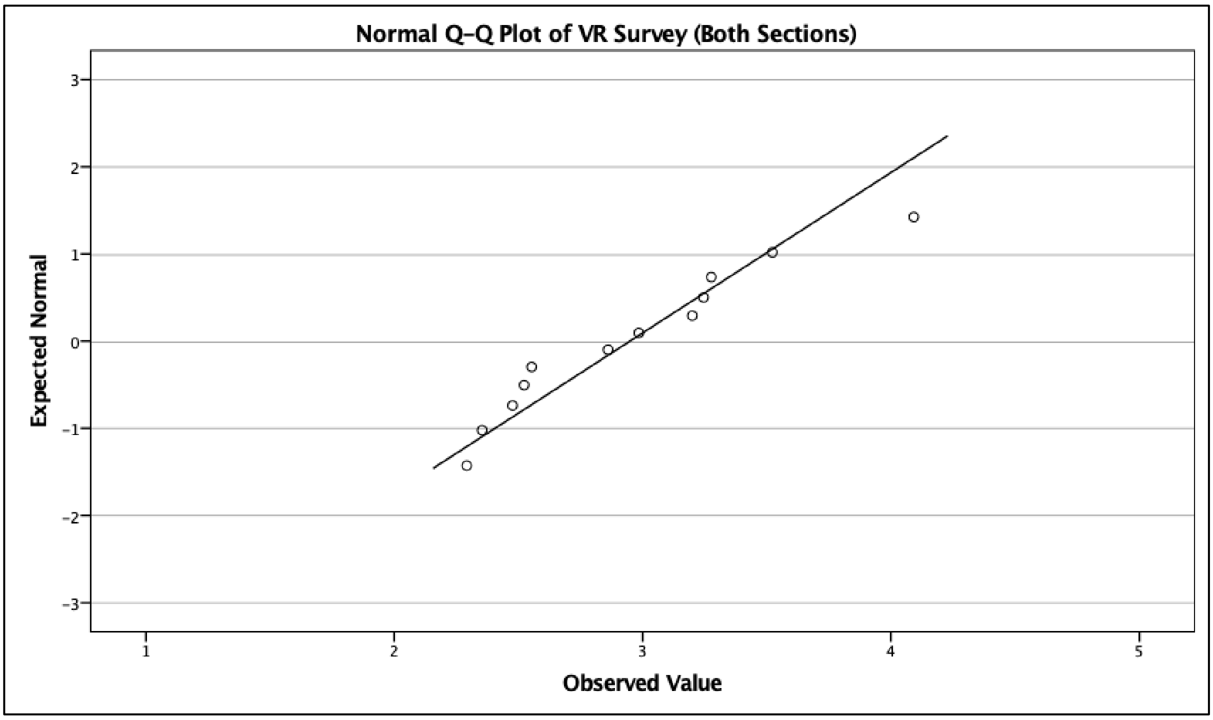


Figure 11: Preliminary Study Game B (VR Version) Q Plot for Complete Survey

The results look slightly skewed, although close to a normal distribution. This can be expected as the sample size is so small (Samuels et al., 2012) (Field et al., 2013) however, the results were further checked by calculating the Z-values for Skewness (S=0.769 / SE=0.637) and Kurtosis (S=1.70 / SE=1.232) both were found to be within the acceptable range of -1.96 and 1.96 for normality (Löfgren n.d.). A Shapiro-Wilk test ($p > 0.05$) (Shapiro & Wilk, 1965) of normality also showed the null hypothesis to retain the 0.05 level of significance, therefore assuming normality for the dataset (Screen $p = 0.102$, VR $p = 0.398$).

Both groups contained the same participants and used the same surveys, with the results normally distributed. This meant a Dependent-means/Paired-Sample T-test was appropriate to compare the means between the playing experience of a screen-based game and the VR game, as it met the necessary assumptions (Field et al., 2013). The results were as follows:

Table 1: Paired Sample Statistics for Preliminary Study Survey (Combined)

Paired Sample Statistics					
		Mean	**N**	**Std. Deviation**	**Std. Error Mean**
Pair	screenboth	2.55000000000	12	.454686907000	.131256804000
	vrboth	2.94871794900	12	.543585345000	.156919573000

The label ‘screenboth’ means the pair contains the results from both sections one and section two of the mobile game questionnaire, whereas the label ‘vrboth’ contains the results from sections one and two of the VR game questionnaires.

Table 2: Paired Samples Correlations for Preliminary Study Survey (Combined)

Paired Samples Correlations				
		N	**Correlation**	**Sig.**
Pair	screenboth & vrboth	12	.675	.016

Table 3: Paired Samples Test for Preliminary Study Survey (Combined)

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	**Std. Deviation**	**Std. Error Mean**	**Lower**	**Upper**	**t**	**df**	**Sig. (2-tailed)**
Pair	screenboth - vrboth	-.398717949000	.410379498000	.118466357000	-.659460642000	-.137975255000	-3.366	11	0.006

The results show a statistically significant result of 0.016, meaning the difference is not due to chance. This is surprising, given the small sample size (Sauro n.d.). Cohen’s d was used to calculate the effect size (Wiseheart n.d.) (Cohen, 1988). This means that, on average, participants had a significantly greater player experience and immersive experience with the VR version of the game when compared to the screen-based game ($M = -0.399$, $SE = 0.410$, $t(11) = -3.366$, $p < 0.05$, with an effect size $d = 1.246$). However, given that the full survey contains questions relating to both the playing experience with respect to the game design and the player’s sense of immersion, it does not give a very detailed picture. We can gain more understanding by separating out the two parts of the survey and comparing each group.

Data Analysis: Section One

With respect to the analysis of the 11 questions written by this researcher regarding the design aspects of the playing experience, the distribution was checked for normality as previously described. Results are reported below:

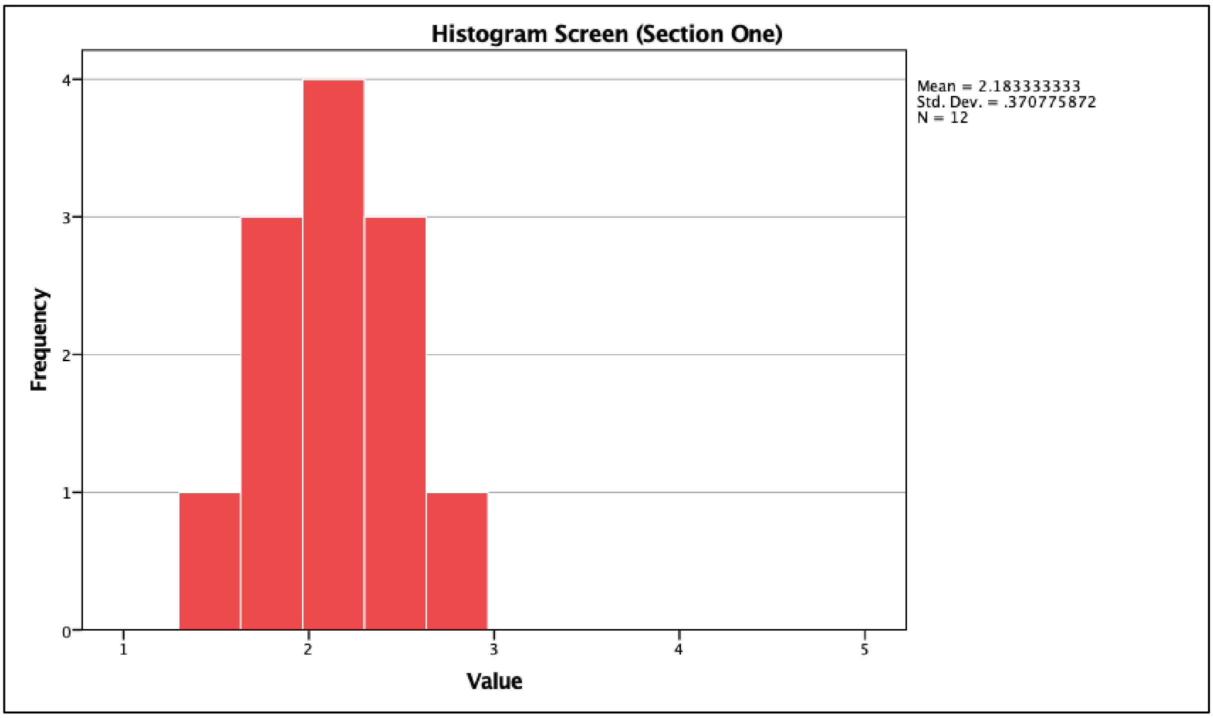


Figure 12: Preliminary Study Game A (Screen Version) Histogram Results for Section One of the Survey

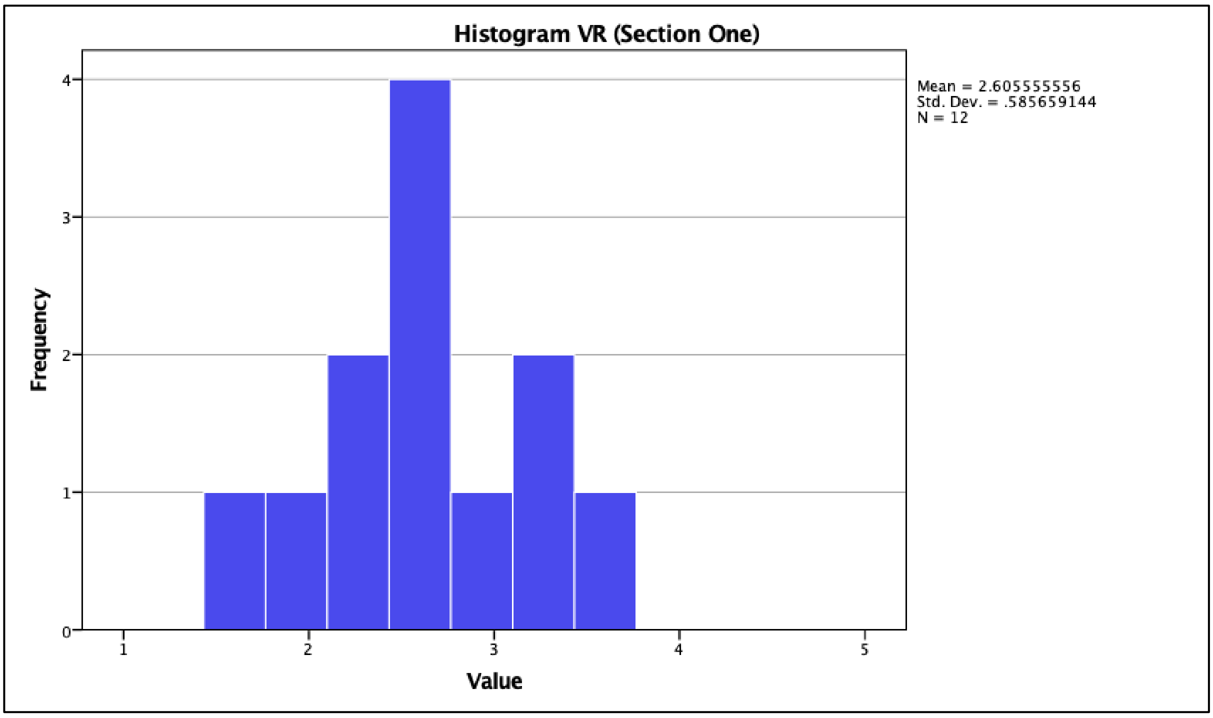


Figure 13: Preliminary Study Game B (VR Version) Histogram Results for Section One of the Survey

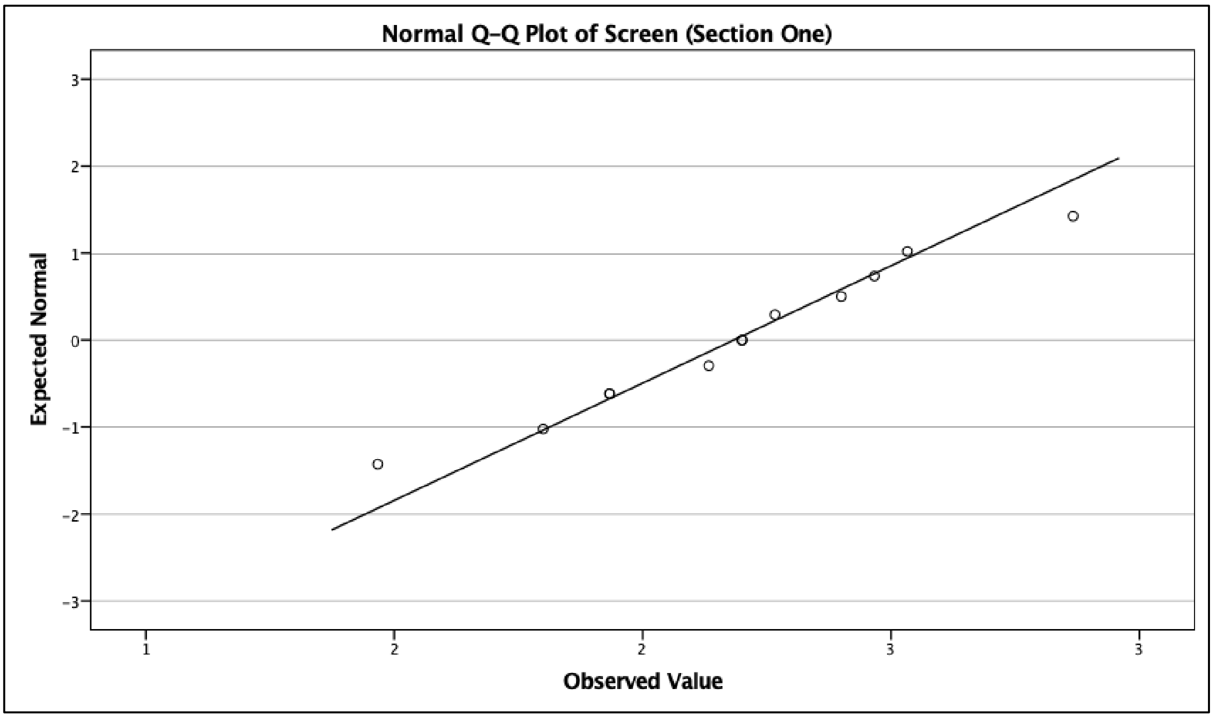


Figure 14: Preliminary Study Game A (Screen Version) Q Plot for Section One of the Survey

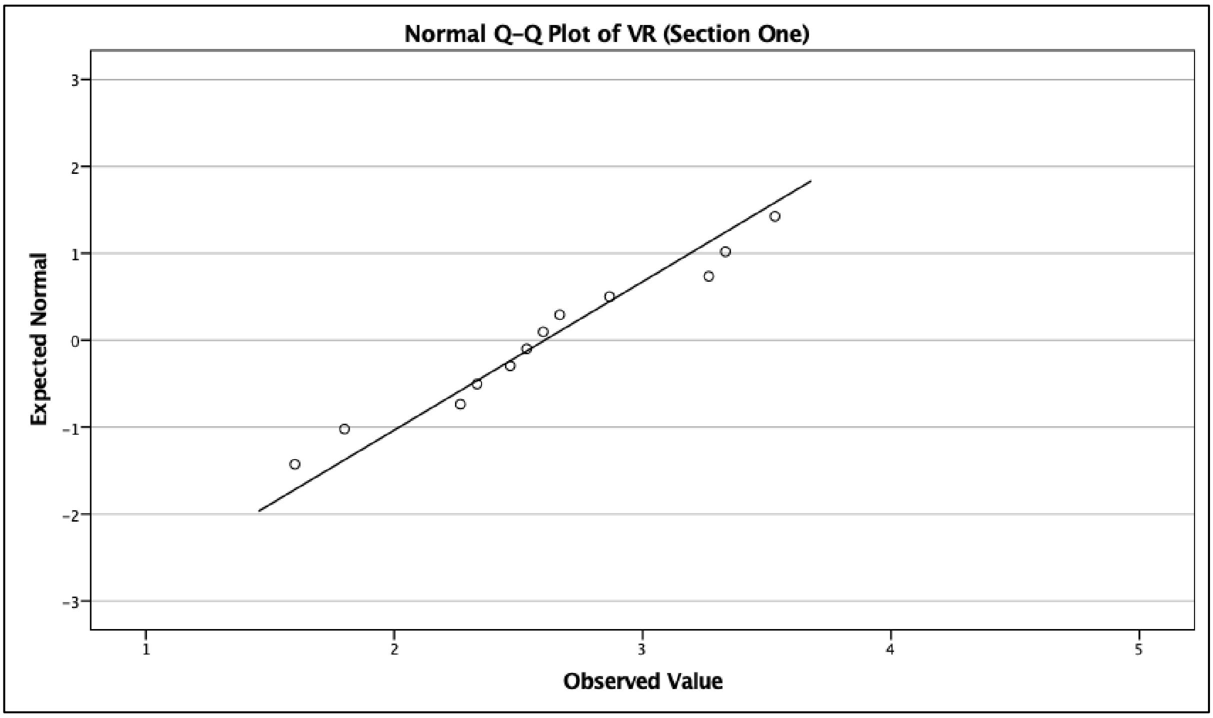


Figure 15: Preliminary Study Game B (VR Version) Q Plot for Section One of the Survey

The distribution of the section one results, as approximated by the eye, looks more normal and less skewed than those of the overall survey results, although we see the values concentrated around a comparatively higher mean of 2.606 for the VR version of the game, compared to a mean of 2.181 for the screen-based version. After further checking the z-values for Skewness (S = -0.129 / SE = 0.637) and Kurtosis (S = 0.427 / SE = 1.232), this was confirmed and found to be in an acceptable range for normality (Löfgren n.d.). The Shapiro-Wilk test also showed the null hypothesis to retain the 0.05 level of significance, therefore assuming normality (screen p = 0.999, VR p = 0.852).

A Paired-Samples T-test was used to compare means:

Table 4: Paired Sample Statistics for Preliminary Study Survey (Section One)

Paired Sample Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair	screenmine	2.18333333300	12	.370775872000	.107033775000
	vrmine	2.60555555600	12	.585659144000	.169065232000

The label 'screenmine' means the pair contains the results for only section one of the mobile game questionnaire, whereas the label 'vrmine' contains the results for only section one of the VR questionnaire.

Table 5: Paired Samples Correlations for Preliminary Study Survey (Section One)

Paired Samples Correlations				
		N	Correlation	Sig.
Pair	screenmine & vrmine	12	0.320	0.310

Table 6: Paired Samples Test for Preliminary Study Survey (Section One)

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair	Screenmine - vrmine	-.422222222000	.584191288000	.168641499000	-.793399659000	-.051044786000	-2.504	11	0.29

Unlike the overall survey results, section one did not show statistical significance, meaning the difference could be due to chance. This is not surprising considering the small sample size (Sauro n.d.). The effect size was also calculated using Cohen's d (Wiseheart n.d.) (Cohen, 1988). Subsequently, on average, participants had no significantly greater playing experience with the VR version of the game compared to the screen-based game ($M = -0.422$, $SE = 0.584$, $t(11) = -2.504$, $p > 0.05$, with an effect size $d = -1.096$).

Given that the small sample size could be affecting the normality of the data, it would be prudent to also compare the means using a non-parametric test for increased sensitivity (Frost, n.d.). The most appropriate test for the sample size is the Wilcoxon Signed-Ranks Test, as there are paired samples (Field et al., 2013). The Wilcoxon Signed-Ranks Test indicated that the VR group scores were statistically significantly higher than the screen-based group scores, with $Z = -2.040$, $p < 0.041$ (Field et al., 2013) (Plummer n.d.). Counter to the Paired-Sample t-test; the Wilcoxon Signed-Ranks does show the participants had a greater playing experience with the VR version of the game compared to the Screen-based version. This is potentially important given the data is ordinal.

The results of the multiple-choice question from the first section are also useful to inform our understanding of the overall player experience. In the VR version of the game, 41.7% of players reported that their sight was the most useful sensory feedback in navigating the in-game challenge, compared to only 8.3% for the screen-based version. Players reported a higher level of enjoyment for the VR version (M = 4.333, SD = 0.778) compared to the screen-based version (M = 3.753, SD = 0.754), finding that the game objective was clearer in the VR version (M = 3.417, SD = 1.165) than in the screen-based version (M = 2.833, SD = 0.937). This was also reflected in how challenging they found each version of the game, where they found the VR was less challenging (M = 3.167, SD = 1.030) compared to the screen-based version (M = 4.333, SD = 0.779). Players also felt significantly more in control playing the VR version (M = 3.833, SD = 1.267) than the screen-based version (M = 2.083, SD = 0.996). Players understood the impact of their sonic expression on their score to the same extent between both versions (VR (M = 2.917, SD = 1.505), Screen (M = 2.917, SD = 1.443) but were more aware of how it changed the accompanying backing track in the VR version (M = 4.417, SD = 0.669) compared to the screen-based version (M = 3.917, SD = 1.240). Players also reported they frequently lost their position within the loop at the same rate between both versions of the game (Screen M = 2.417, SD = 0.900, VR M = 2.417, SD = 0.996). Evaluating these questions and answers in relation to the demographic may also prove useful in furthering understanding of the overall playing experience. This will be done after analysis of section two of the survey.

Data Analysis: Section Two

The distribution for the remaining 44 questions relating to player immersion was checked for normality as previously described. The analysis is reported below:

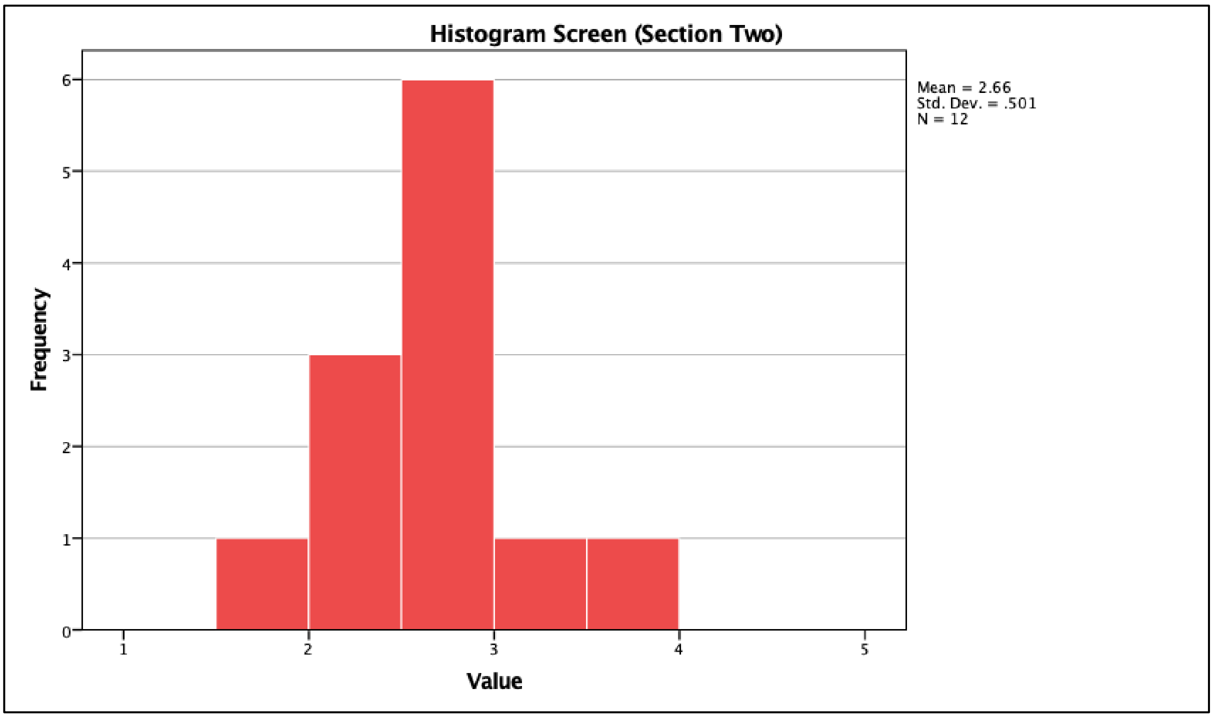


Figure 16: Preliminary Study Game A (Screen Version) Histogram Results for Section Two of the Survey

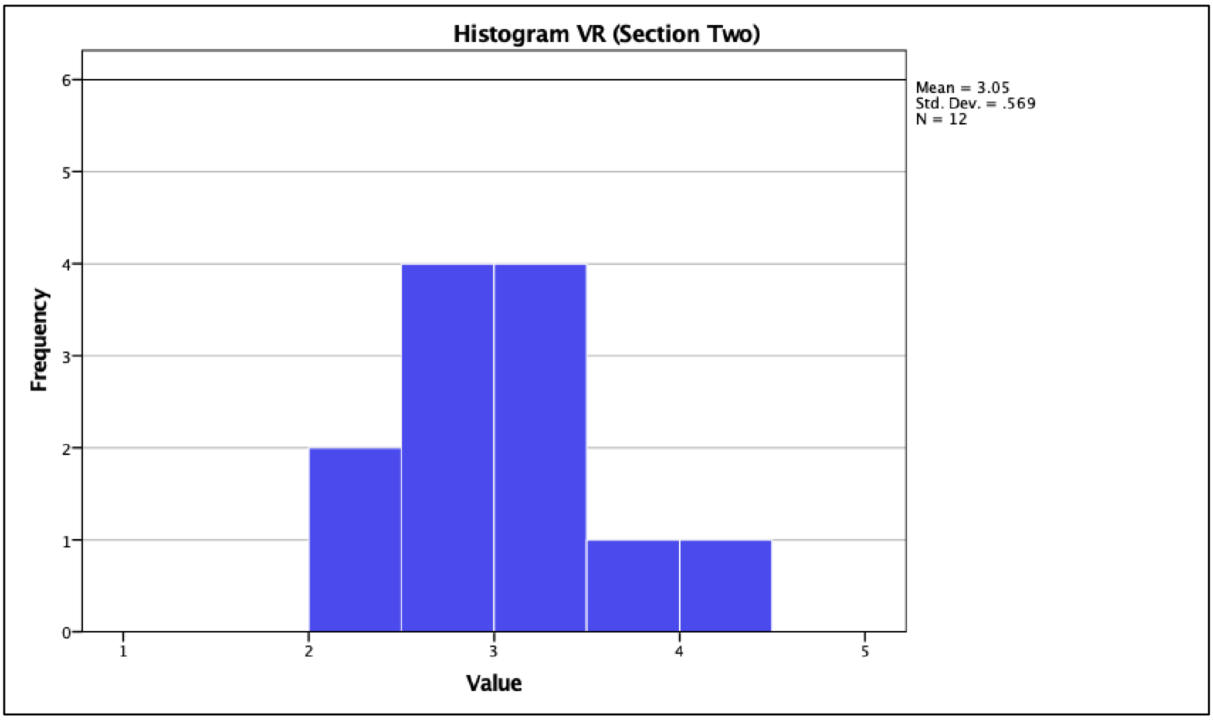


Figure 17: Preliminary Study Game B (VR Version) Histogram Results for Section Two of the Survey

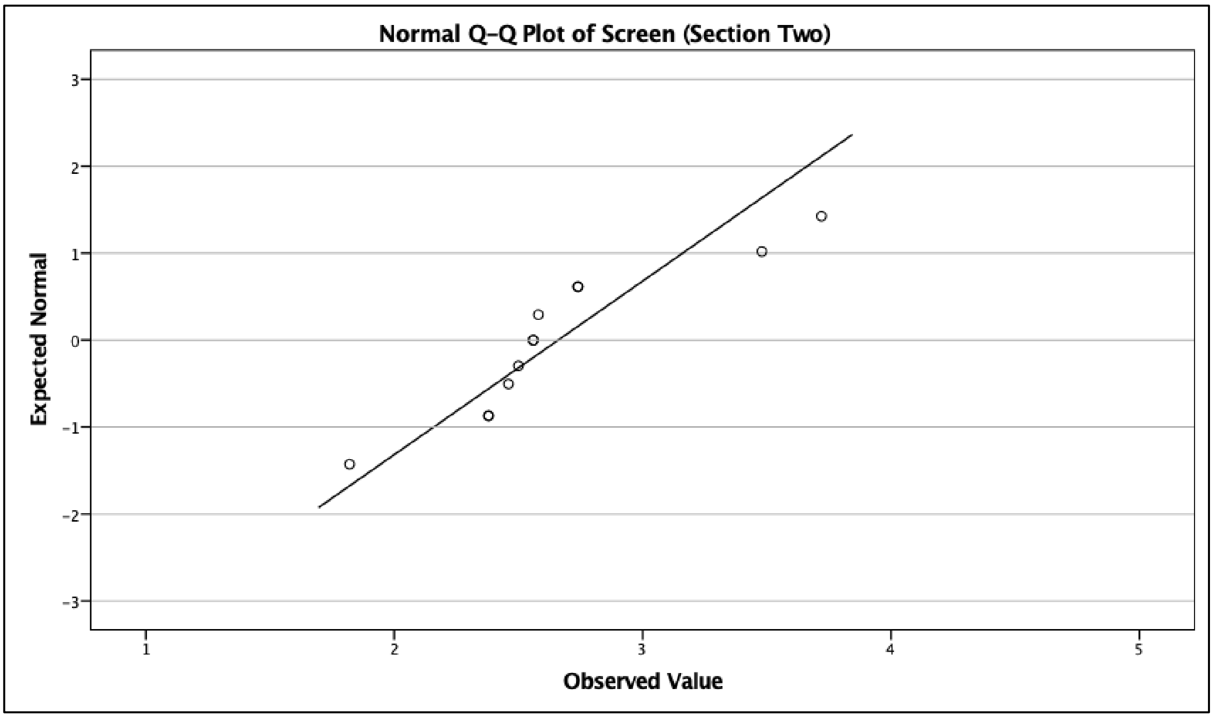


Figure 18: Preliminary Study Game A (Screen Version) Q Plot for Section Two of the Survey

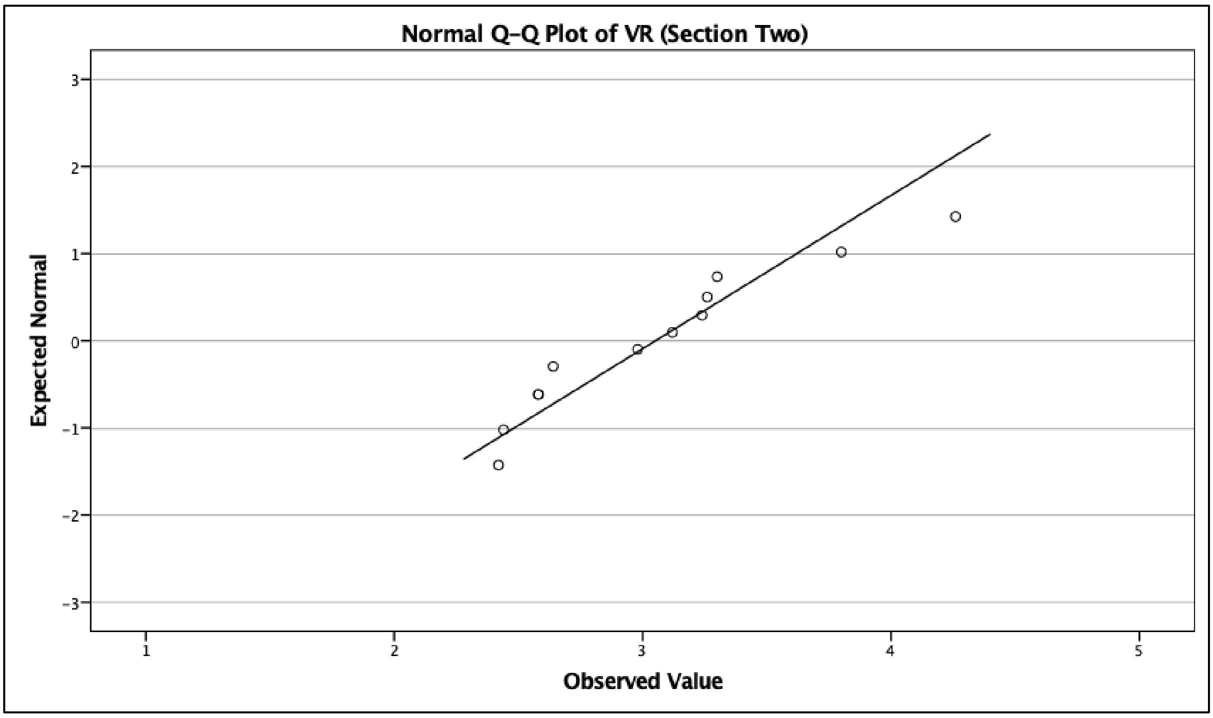


Figure 19: Preliminary Study Game B (VR Version) Q Plot for Section Two of the Survey

The distribution of results for section two is approximated to be normally distributed, although they show some skew like the overall survey distribution. After checking the z-value for Skewness (S = 0.946 / SE = 0.637) and Kurtosis (S = 1.529 / SE = 1.232), they were found to be within the acceptable range (Löfgren n.d.). The Shapiro-Wilk test also showed the null hypothesis to retain the 0.05 level of significance, therefore assuming normality (screen p = 0.05 VR p = 0.194).

A Paired-Samples T-test was once again used to compare means.

Table 7: Paired Sample Statistics for Preliminary Study Survey (Section Two)

Paired Sample Statistics					
		Mean	**N**	**Std. Deviation**	**Std. Error Mean**
Pair	screenpart2	2.6600	12	.50120	.14468
	vrpart2	3.0517	12	.56873	.16418

The label 'screenpart2' means the pair contains the results for only section two of the mobile game questionnaire, whereas the label 'vrpart2' contains the results for only section two of VR questionnaire.

Table 8: Paired Samples Correlations for Preliminary Study Survey (Section Two)

Paired Samples Correlations				
		N	**Correlation**	**Sig.**
Pair	screenpart2 & vrpart2	12	.768	.003

Table 9: Paired Samples Test for Preliminary Study Survey (Section Two)

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	**Std. Deviation**	**Std. Error Mean**	**Lower**	**Upper**	**t**	**df**	**Sig. (2-tailed)**
Pair	screenpart2 - vrpart2	-.39167	.36952	.10667	-.62645	-.15689	-3.672	11	.004

The results show a statistically significant result of 0.003, meaning the difference is not due to chance. The effect size was also calculated using Cohen's d (Wiseheart n.d.) (Cohen, 1988). Subsequently, on average, participants had a significantly greater playing and immersive experience with the VR version of the game compared to the screen-based game (M = -0.3917, SE = 0.370, t(11) = 3.672, p < 0.05, with an effect size d = -1.075). Consequently, on average, participants had a greater sense of immersion in the VR version of the game than the screen-based game. This section is best analysed as a whole. However, there are some suitable questions that can be individually observed to give further insight into the playing experience. Players reported putting less effort into the VR version (M = 4.083, SD = 0.669) compared to the screen-based version (M = 4.25, SD = 1.138). Players reported that they felt like they were interacting with the game environment more in the VR version (M = 4.167, SD = 0.937) than the screen-based version (M = 3.5, SD = 1.087). Players felt like they were moving through the game according to their own will in VR (M = 3.583, SD = 1.311) significantly more than the screen-based version (M = 2.5, SD = 1.1677). Players thought they performed better in VR (M = 2.833, SD = 1.267) compared to playing the screen-based version (M = 1.917, SD = 0.996). These support the findings from the first section of the survey.

Data Analysis: IEQ 31 Questions

The results from the original 31 IEQ questions (31 five-point Likert; the additional 10-point Likert baseline question was not asked) are documented herein. The IEQ provides scores for five immersion factors; 'Cognitive Involvement', 'Emotional Involvement', Real-World Dissociation', 'Control', and 'Challenge' (Jennett et al., 2008). Jennett uses the total of these 31 questions to provide a score for immersion. This score ranges between 31 and 155, where the 'score could be used as an indicator whether participants are engaged (between 31- 90), engrossed (between 90-120) or are totally immersed (between 120- 155)' (Nordin et al., 2014) (Jennett et al., 2008). As the five factors are comprised of a different number of questions per factor (nine of the 31 questions determine the factor of 'Cognitive Involvement', six determine the factor of 'Emotional Involvement', seven determine for 'Real-World Dissociation', five determine factor of 'Control', and four are used to determine the factor of 'Challenge'), and I am interested in comparing individual factors (the motivation explained within the evaluation later in this chapter) the totals for each factor have been normalised to the one-to-five five-point Likert scale for clarity. This allows for direct comparison between factors as they are all appropriately scaled. Multiplying the mean of the factor by the corresponding number of questions (for example, multiplying the mean of 3.870 for the factor of 'Cognitive Involvement' in game A (as found in Table 10) by nine) will give the score for that factor should a comparison between scores be necessary. The following table contains the breakdown per immersion factor for the screen version of the game.

Table 10: IEQ Scoring Immersion Factors for the Screen-based Game (Preliminary Study)

IEQ Scoring Immersion Factors						
	Cognitive involvement	**Emotional Involvement**	**Real World Dissociation**	**Control**	**Challenge**	**Overall**
Mean	3.630	2.986	3.155	3.05	3.646	3.306
Std. Deviation	0.671	0.920	0.722	0.427	0.470	0.477

The breakdown per immersion factor for the VR version of the game is below.

Table 11: IEQ Scoring Immersion Factors for the VR Game (Preliminary Study)

IEQ Scoring Immersion Factors						
	Cognitive involvement	**Emotional Involvement**	**Real World Dissociation**	**Control**	**Challenge**	**Overall**
Mean	3.870	3.278	3.917	3.7	3.25	3.658
Std. Deviation	0.678	0.770	0.595	0.721	0.440	0.492

We can see in the table above that for four of the five factors, players reported an increase in immersion in the VR version of the game (VR M = 3.658, screen M = 3.306). This is comparable to the results found by Thompson concerning the effect of touch-screen size on game immersion, where players reported that a larger screen resulted in higher levels of immersion; an iPad (M = 3.761) compared to an iPod Touch (M = 3.3) (Thompson et al., 2012). The only factor that was higher in the screen version of the game was for 'Challenge'. We can also see the combined highest reported factor across both versions of the game was 'Cognitive involvement'. This would speak to the player's reported engagement with the in-game task. However, for both versions of the game, there was a higher individual factor. In the screen version, the results show 'Challenge' as the highest immersion factor (M = 3.646); this would reflect how difficult players found the interaction method in the game. However, it is worth noting this did not create a negative impact on their sense of immersion, as it was the only immersion factor that was higher for the screen version compared to the VR version of the game. For the VR version of the game, 'Real World Dissociation' was the highest immersion factor (M = 3.917). Given the enclosed nature of the VR headset, it physically shuts out everything else from the player's periphery vision, which is not surprising. This argument can be used to justify the difference between the two versions. (Screen M = 3.155, VR M = 3.917). Players reported significantly higher for the VR version of the game for the immersion factor of 'Control' (M = 3.7) compared to the screen version (M = 3.05) of the game. This is likely due to the difference in quality of the embodied interaction for placing the cubes and triggering the sound. The difference in the remaining factor of 'Emotional Involvement' (VR M = 3.278, Screen M = 2.986) could also relate to the quality of the embodied interaction relative to the affective nature of music creation. A player in more control will likely feel more responsible for their actions, likely impacting their emotional experience (Juul, 2009).

Demographics

The most important demographic questions asked related to the prior gaming experience and musical experience of the participants. All 12 participants had played video games for over five years. 40% of the participants had played a couple of times per month, with 20% playing five or more days per week, with another 20% playing every day. Only one participant reported they played with a VR headset. The most popular platform was a console without motion controllers, followed by a PC without motion controllers. Of the nine participants who had ever used motion controls, the majority (55.6%) played games that utilised movement a couple of times per year. These participants with motion control experience found the game objective clearer (M = 3.333, SD = 1.073) than those without (M = 2.917, SD = 1.084), found the interaction more intuitive (M = 4.167, SD = 0.718) than those without (M = 3.417, SD = 0.793), found the sound input more responsive (M = 4.333, SD = 0.651) than those without (M = 3, SD = 1.206), and found the visual feedback of the game more useful (M = 3.417, SD = 1.084) than those without (M = 2.833, SD = 1.193). 75% of participants had experience playing a musical instrument, with the majority (55.5%) of them playing for at least five years. A third of these practised a couple of times a month, with another third practising a couple of times per year. 58.3% of the participants had experience using music production software. All these participants were still active as digital music practitioners. Participants with experience of using music production software found the game objective clearer (M = 3.5, SD = 1.091) than those without (M = 2.6, SD = 0.843), felt more in control (M = 3.214, SD = 1.577) than those without (M = 2.6, SD = 1.174), and clearly understood the effect the note placement had on their score (M = 3.143, SD = 1.351) compared to those that didn't (M = 2.6, SD = 1.578).

Evaluation

For evaluating the data analysed, it is worth tempering expectations of what can be drawn from these results as the sample size was small. However, it indicates potential trends worth exploring further. The results between the two different groups in the full post-survey do indicate that the playing experience changed significantly, despite being vague as to how exactly the playing experience has changed. Analysis of section one of the post-survey suggests there is potential for music games to be better experienced in VR compared to the traditional screen approach, although it is important to consider that the questions in this subsection of the survey are very broad. It is also worth considering that many participants found the game very difficult, as the in-game challenge was highly abstract and paired down. Subsequently, this may be impacting the results. Section two of the survey shows that the participants found the VR version to be a more immersive experience. This is not surprising considering the nature of VR: wearing a headset, which situates oneself in the in-game spatial and sensory environment, isolating oneself from the real-world surroundings.

The scoring of the IEQ supported the analysis from section two and provided a more detailed breakdown. The VR version of the game was found to be more immersive in four out of the five immersion factors of the IEQ. These were: 'Cognitive Involvement', 'Emotional Involvement', 'Real World Dissociation', and 'Control'. It is not surprising that the factor of 'Real-World Dissociation' (VR M = 3.917, Screen M = 3.155) was higher, given the nature of wearing a VR headset. However, the remaining three factors are interesting to evaluate further. The difference in 'Cognitive Involvement' (VR M = 3.870, Screen M = 3.630) is curious as both versions of the game required the same task. This difference could be due to the additional demand on spatial reasoning when moving around in VR, as players had the opportunity in that version of the game to move the camera around the visual elements. It was possible for them to view the 3D graphics from different positions, which may have encouraged some more cognitive involvement in trying to 'solve' their next expression. The difference in 'Emotional Involvement' (VR M = 3.278, Screen M = 2.986) is also fascinating. Is this a by-product of a more meaningful interaction? (McMahan, 2003). Is this a benefit of a higher technological affordance (Gibson, 1979) (Gaver, 1991) leading to a more immersive, embodied interaction that is more affective in nature? The difference in the factor of 'Control' (VR = 3.7, Screen M = 3.05) is also supported by the findings of section one. Players found the interaction significantly more intuitive in the VR version. By not having to focus their attention (Cairns et al., 2014) (Jennett et al., 2008) (Grodal, 2003) (Klimmt, 2003) on what they were physically doing to make their input, players had a more immersive experience. This would support the observation regarding technological affordance that was made for 'Emotional Involvement'. By not having to think about their interaction directly, players could focus their attention more on their desired musical outcome. The only immersion factor that was higher in the screen-based version of the game was 'Challenge'. Players reported they found the game much more challenging when playing the screen-based version compared to the VR version. This was likely due to the difficulty they experienced regarding the interaction method. Even though players found the challenge very difficult, it did not appear to put them off from engaging with the game. However, as found in the results of section one, players reported a more enjoyable playing experience in VR (VR M = 4, Screen M = 3.583), which potentially highlights the importance of a well-designed challenge.

The study also does not factor in the multitude of different interface possibilities available for both screen and VR gaming platforms. Although the selected interaction method for the screen-based game was chosen explicitly as the action of swiping, as an easy input method, many participants found it difficult. Consequently, the differing results relating to their experience regarding interaction may be non-typical. Despite these difficulties, players did still want to 'win' the game and wanted to continue playing. They also considered both to be reflective of a musical experience (screen M = 3.71, VR M = 4 Q14 section one), favouring the VR version. Coupled with the difficulties players had with the screen version, this is interesting. It is further supported by considering that the players, on average, found the controls more intuitive in the VR version (screen M = 3.35, VR M = 4 Q6 section one). It was also evident through the breakdown of demographic characteristics that players' prior experience with motion controls benefitted the player's understanding of the objective (with experience M = 3.333, without experience M = 2.197) and impacted their perception of how responsive the sound interaction felt (with experience M = 4.333, without experience M = 3). Does this mean the quality of gestural interaction plays a part in how immersive the experience can be? To what extent are intuition, sonic expectation and sensory-motor processes used in inducing immersion during play/performance? These questions and how they relate to each other warrant further study. Gestural interaction, intuition, and sonic expectation fall within the scope of my research, although in a limited capacity with respect to game design. Sensory-motor processes and low-level sound cognition, whilst very interesting, are outside the remit of this research. The quality of non-auditory feedback is also important. For example, how does visual perception and the use of stereoscopy in VR create the embodied interaction? And to what extent does this impact the playing experience? These are part of this research as they influence the player's perception of their interaction, although they are not the primary focus.

Regarding the difficulties caused by interaction, how does the difficulty of the challenge impact the playing experience? Does a less explicit process to complete the in-game challenge result in a negative player experience? Would the playing experience differ if the challenge were still as severe or implied but related to a non-musical context? Is this a unique issue that relates to musical expression, and does existing musical performance-related research support this argument? Are there parallels between the gaming playing experience and musical performance, and with what parity? Given that both playing games and playing a musical instrument take a lot of time to learn (and it is only after players gain some skill in both they can feel comfortable enough to experiment), how can these issues regarding the challenge of play be rectified?

The findings from the demographic breakdown highlight that prior musical experience, in relation to the use of music production software, did have an impact on how the players found the interaction. Those with experience felt significantly more in control (M = 3.214) than those who lacked experience (M = 2.6). This potentially highlights the relationship between sound and challenge and identifies how sensitive players are to the contextualisation of the challenge when it involves music. Players could be given more time before evaluating their experience, or it could be evaluated at different intervals to see if their playing experience improves over time. Would altering the challenge of the game so it is more game-like and less like an instrument provide this comfort? If the players are rewarded earlier in the learning curve of the game, would they feel like they are having a more positive experience and impact on the game? These are all design questions, problems, and factors to consider for the next prototypes. The most important three are ‘Does a less explicit process to complete the in-game challenge result in a negative player experience?’, ‘How does the difficulty of the game challenge impact the playing experience regarding difficulties caused by interaction?’ and ‘How can the challenge of play be rectified?’

There is also the potential issue relating to the player’s expectations of the game, making it difficult to present musical ideas as players are expecting an explicit challenge. Players also get self-conscious expressing themselves, especially when they are learning a new tool in front of other people (in front of the researcher). This is particularly true if they have already had experience playing a musical instrument and have given up. Regardless of the reason, when the challenge is self-led, it is a double failure of both the performance of the instrument and the performance of generating a challenge at which they can succeed. Moving expression into part of a challenge, or the tool for a challenge, and not the challenge itself, may negate this.

However, the Preliminary Study is enough to point to a potential trend and certainly can be used as a basis for further studies. Narrowing down the range and focus of questions in Part 1 to a smaller topic would be beneficial. Assessing the quality of interaction and the musical experience in more detail would most likely be an improvement. This could then be tested by the addition of a control group, with the game minimising the player’s agency over their musical expression. For example, asking the player to explicitly pattern match to score as opposed to scoring continuously based on their previous input.

Conclusion

VR lends itself to a comparatively more immersive experience than that of a traditional screen. Given the aim of this research is to use musical expression to aid player immersion, it is imperative to work primarily in a medium conducive to player immersion. Therefore, future prototypes will be built for VR, with the following studies focusing on ways to measure the experience in greater detail. It is essential that the in-game challenge in the game is not contextualised as a music challenge, as this potentially imposes an additional challenge of comprehension. Subsequently, it is important that in-game challenges are designed in a way to mitigate any potential lack of musical experience, as this makes it more accessible for players.

Study Two: The Impact of Player Agency Over Rhythmic Sequencing and Platform Placement on the Playing Experience

Abstract

This chapter presents a study of player experience in an experimental VR platform game. In the game, players are given a free choice of note placement and corresponding platform placement compared to a predetermined sequence designed by this researcher. This takes place in a prototype 3D VR platform game called 'Red is Dead'. It explores the impact of the player's creative decisions to navigate a typical in-game challenge of a 3D platformer on their enjoyment and their sense of immersion. The results of this study will help to elucidate the role of agency in music-game design and whether it can be leveraged to make engaging novel game experiences. It is hoped that the outcome of the study will establish to what extent players consider their musical-expressive decisions as opposed to their desire to complete an in-game challenge and the impact this has on their playing experience.

Introduction

This study was conducted to understand the importance of player agency in musically expressive in-game decisions. The study was a 'within group' study with two different versions of the game in an A/B testing format. A total of 30 participants took part and played both versions of the game. The study focused on the exploration of differences in player experience and the player's sense of immersion in the specific context of a VR 3D platform music game when exposed to differing degrees of agency. The goal of the game was to make use of sonically expressive musical interaction to navigate from a starting point to an endpoint in a 3D environment. The rationale behind the study is to inform future game design. Regardless of the outcome, further research will be based on the resulting outcomes of this study. If player experience differs greatly between the two different versions, then further research will aim to study in detail what could cause such different outcomes. If there is no difference in playing experience, then further research will attempt to address what range of expression players would consider necessary to make an impact on their experience and sense of immersion. The outcome of this study will improve our understanding of what impacts a player's sense of immersion, including whether player agency can help or hinder immersion and to what extent the player's musical expression impacts their playing experience. It will also improve our understanding of whether agency in expressivity is relational to the player's attention on completing the gaming challenge. It was decided at the outset that, following this study, further research will be narrower in focus based on these findings.

This has important implications for game development regarding challenge and level design. If it proves beneficial for players to be given more autonomy over their sonic experience, then further work should encompass and document this in greater detail. This would increase the overhead and complexity of future games, as providing multiple different potential outcomes or uses by the player requires considerably more design, planning, and computational logic than a single predetermined outcome. It would also subsequently require more testing to evaluate final designs. Outside of the value of a potential method for a game to increase player engagement, it is also worth considering that for any-sized game studio, this would come with both an economic cost and the cost of time spent. Before a cost-benefit analysis could be conducted on whether the inclusion of such an approach would be worthwhile to consider, it would be prudent to determine to what extent it is effective.

Methodology

A 3D platform game was developed to provide a typical gaming challenge. The game contained features designed to test sonic expression and rhythmical arrangement, encompassed as a gaming mechanic integral to the overall game design. The first feature is a simple interactive method for the player to generate sounds, which includes the ability for the player to arrange when the sounds should be triggered within a set repeating periodicity. The second feature is a challenge for them to complete. The first feature is necessary as without the ability to generate and place different sounds, we would not be able to answer our question. The second feature is necessary for contextual framing. There were two different versions of the game used in the study. All participants played both versions in a random order. The 'A' version allowed the player to select and place any one of four different percussive sounds. The 'B' version did not give the player any agency over what sound or where the sound can be placed. The game was developed using JavaScript with the MaxiInstruments library for audio events and the A-frame library for 3D graphics. This was chosen over the previous use of Unity from the Preliminary Study as a response to the COVID-19 outbreak. Developing specifically for a low overhead web platform on a commercially available headset (Quest 2) allows for easier in-browser delivery and surveying, as opposed to a set location with an Oculus Rift Dev kit and specific PC.

Game Design and Dynamics

The game introduces the broad concept of musical rhythm and arrangement. The challenge of the game is to navigate from a starting position to an end position that exists on the y-axis directly above the player. This end platform is visible to the player if they look up and it follows them around the level. The player is instructed to arrange sound to reach this endpoint while avoiding enemy 'playheads'. To arrange sound the player must collect tokens located around the level. In order to navigate to the tokens, the player must walk/run and jump along a series of platforms. This is accomplished by using the analogue sticks on the controllers for player movement and the facing direction of the headset as the direction of the player's camera. The player starts the game with a platform in front of them along the z-axis, with a red playhead moving above the platform in a cycle (see Figure 20 below). For each 'row' of platforms, there is a token. To generate sound and further platforms, the player must collect these tokens. Each row of platforms is comprised of 16 individual tiles that correspond with a step in a step sequencer. In the 'A' version once a token is collected of the game, the player can press any one of four buttons to place any one of four samples, in the current corresponding step the player is standing over on the platform. Each platform is comprised of 16 tiles, with each relating to a corresponding step of the sequencer. When new platforms are created the starting position of these tiles are rotated, this keeps the playheads for each platform in sync (further explanation can be found in the game design document in the appendix on page 237). If the player is unhappy with any of their position notes, they can press a button to delete that specific input. The player will recover that token to use elsewhere; however, as a mild punishment for poor placement, the playhead from the original input will remain cycling through its movement in the game environment. Subsequently, the player must be mindful in placing tokens. In the 'B' version of the game, the player is not given this choice – Once the player collects the token it generates a note and platform in that position. The control scheme for both versions of the game can be found in the table below.

Table 12: Player Controls for Study Two Game

Controller Input	Game A	Game B
Thumbsticks (up and down) and direction player is facing	Forward/backward movement	Forward/backward movement
Thumbsticks (left and right)	Strafing	Strafing
Left Trigger Button	Kick drum note	N/A
Right Trigger Button	Snare drum note	N/A
Left Grip Trigger Button	Hi-hat 'bleep' note	N/A
Right Grip Trigger Button	Cowbell 'bloop' note	N/A
X or A Button	Jump	Jump
Y or B Button	Destroy note/step value and platform the player is currently standing on	N/A

At the same time as the sound generation, it also generates a new row of platforms from that starting position, one unit up in the y-axis, in the direction the player is facing. Each of the four sounds has a corresponding platform form (colour, width, and height), different from the other three sounds/platform types, along with a different playhead form. In Figure 22 we can see a token has been placed on top of the white starting platform. The width size (3x3) and orange colour mean this token was converted into a kick drum, in that specific step of the sequencer. Additionally, according to where each note has already been placed, that specific tile will be shaped relative to the type of sound in place. For example, the kick drum sample is a 3x1x3 tile; if there is a kick in steps 1, 5, 9, and 13, then those tiles will appear larger in width and depth.

By arranging the sound, the player creates new platforms to stand and jump, allowing themselves to navigate around the game environment, eventually moving their way up to the end goal platform. If the player is knocked off by the playhead or falls off the platforms, the player falls for a short duration and hits a red kill plane. The game then places the player back at the starting position. In the 'A' version of the game, the coherency of placement impacts where the tokens are placed around the game environment. For simplicity's sake, we are defining coherence to be a moderate distribution of each sound. If the player tries to *game* the mechanic by placing the notes in the same place or in the position of the collected token, the tokens are further away. Additionally, the enemy playhead for the generated platform will increase its movement speed, making it harder to avoid. In the 'B' version, they are in a fixed sequence.

In the 'A' version of the game (where the player has agency over where to place a note), a typical playing experience would be the following:

1. The player has been given instructions to collect yellow tokens. They start at the end of one platform and can see a token at the opposite end of the platform. Backing music can be heard. In between the player and token there is a moving red playhead (see figure 20).
2. The player moves towards the token using the analogue stick, they use the 'X' or 'A' button to jump.
3. The player tries to avoid the moving red playhead. If they get hit the sound generates a tone and the player gets knocked off the platform, falling to the red kill plane. As the player falls the backing music has some of its parameters changed (further details on which specific parameters can be found in the game design document in the appendix on page 237). The player is then placed back at the starting position on the first platform.
4. Having avoided the moving red playhead, the player collects the token.
5. The player can now place their choice of four percussive sounds, using any of the four triggers, at any step of the step sequencer. This is accomplished by standing over the corresponding tile of the platform and pressing a button.
6. Once a note has been placed an additional platform appears. This originates from the position of the placed note, albeit one unit higher on the Y-axis. There is also a new token to collect in the level, along with an additional red playhead to avoid. The size and shape of these will be dependent of which of the four sounds/platforms have been generated.

7. The game checks the coherency of note placement relative the previous input. Depending on this coherency, the position of the next token is determined.
8. The player repeats steps 2 to 7 until they can reach the end platform.
9. The player jumps on the end platform. The player is taken to an end game level.

In the 'B' Version of the game (where the player does not have agency over where the notes can be placed), a typical playthrough would be:

1. The player has been given instructions to collect yellow tokens. They start at the end of one platform and can see a token along the platform in front of them.
2. The player moves towards the token using the analogue stick, they use the 'X' or 'A' button to jump.
3. The player tries to avoid the moving red playhead. The outcome of being hit by the playhead is the same as it is in the 'A' version above.
4. Having avoided the moving playhead, the player collects the token. Upon collection it places a predetermined percussive sample/note in the corresponding step of the step sequencer relative to the position of the token over the platform. It also generates a corresponding platform, enemy playhead, and a new (predetermined) token to collect.
5. Steps 2 to 4 are repeated until the full predetermined rhythm has been created and there are enough platforms to reach the end platform.
6. The player is taken to an end game level.

In addition to the percussive sounds collected and placed by the player, there is also backing music that cycles through different arrangements based on the number of tokens picked up. This is generated by three synthesisers: a bass synth, a mid-range synth, and a synth to trigger notes when the player collides with a playhead. These are parameterised, with parameters changing based on the player's behaviour. There are also several different filters and audio effects applied to the audio signal based on the player's input, interaction, and position in the game world. When the player jumps, a low-pass filter is applied to all the percussive samples - each with an independent scale for cut-off. It also changes the LFO rate and LFO pitch modulation for the bass synthesiser and the LFO frequency parameter for the mid-frequency synthesiser. When the player is hit by a playhead, a random note (assigned from a predetermined scale) is triggered, along with a delay and reverb applied. If the player falls past the lowest platform, a hi-pass filter is applied to the master audio signal, along with the gain of each synth being reduced to zero until the player is reset to the starting position.

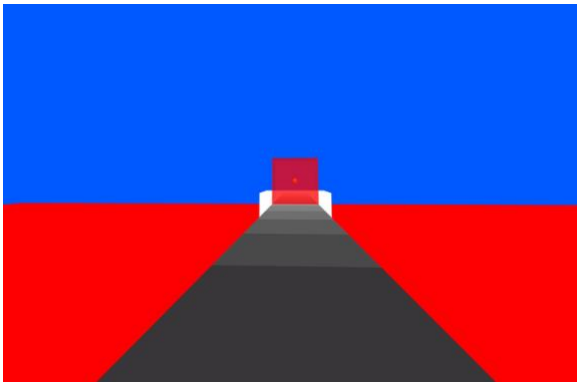


Figure 20: Red is Dead: The player is on the first platform

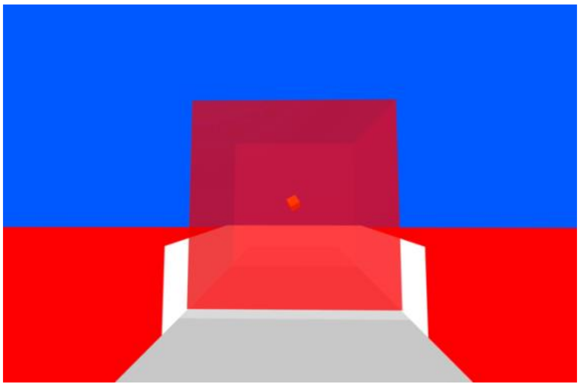


Figure 21: Red is Dead: The playhead is coming towards the player, a token to collect is visible behind it

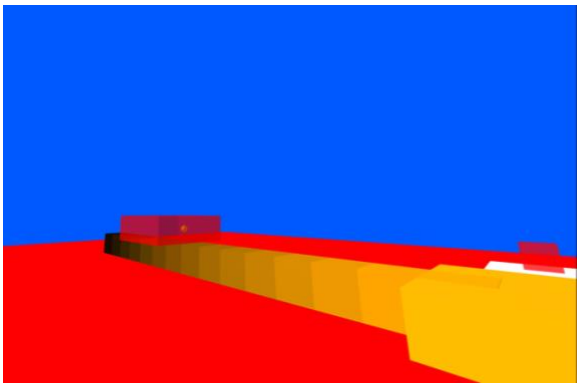


Figure 22: Red is Dead: A token has been collected and placed

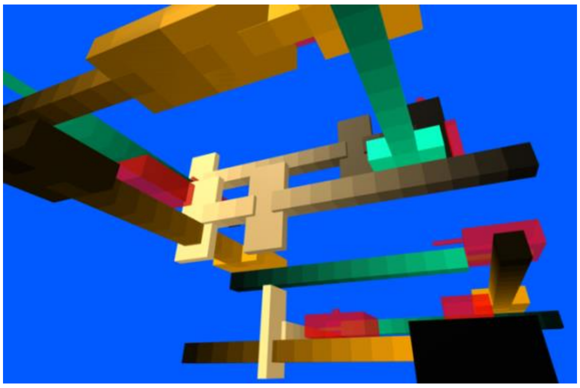


Figure 23: Red is Dead: Many tokens have been placed through the environment!

Although the player has control over the choice of sample and note placement (and to an extent when to move through the backing arrangement), they do not have control over the overall sonic aesthetic. Each section of the arrangement is fixed. The main features of each synth, the wave shape, LFO, filter, frequency type, and, consequently, timbre, are all fixed. These have all been deliberately designed by this researcher. The aggressively haunting electronic drone arrangement, along with deliberately synthetic sounding percussive samples (with specific parameters utilised with player interaction), was an intentional choice. The sound is stylised towards late 80s electronica / early 90s IDM, a style chosen as it accompanied the pared-down graphics of the game well. Although the game is designed specifically to test sonic expression, there were still some considerations made for graphics. The 3D graphics needed to be clear to see and easy to interpret while moving or looking around the environment quickly. Everything red in the game will result in a bad outcome for the player (hence the name of the game). The playheads that knock the player off are red, and the kill plane at the bottom of the level is red. Each sample/platform type is a different colour, with the colour palette being a deliberate attempt at late 80s PC game stylisation.

There is no time limit for how long the player can spend collecting tokens. Due to the different shapes of the platforms and depending on where the player chooses to arrange notes, it is possible for the player to remain stationary and contemplate their work. The game only concludes when the player lands on the end platform. Once this happens, it triggers an end cycle; every four repeating cycles, it removes some of the instrumentation, along with changing the background and platform graphics to black. Upon the final cycle, it restarts the level in a 'free' range mode. This was included as a bonus reward for participants able to complete the game. In the free mode, the player can trigger notes at any point without the need for tokens, leading to some interesting results!

Once the game was developed, it was provisionally tested for bugs and performance-related issues, receiving some optimisation for performance for the Oculus browser. A splash webpage was created detailing necessary game information for the study with a randomised link to the first version of the game, and additionally provided the necessary GDPR for online data collection (Goldsmiths, University of London, 2023). The link to the version of the game was randomised to avoid any potential bias arising out of the playing order of the two versions of the games - there is potential for players to find the game easier or more intuitive once they have gained playing experience. As the aim was to recruit participants online, it was accepted they would likely be using their own headset and subsequently would be familiar with the basic functionality of the device. Consequently, there was no need to introduce them to the device. Participants recruited locally were asked if they had any prior experience with the headset and given some time to gain familiarity with the controls before starting. As the in-game player movement was controlled via the analogue sticks and head position, there was no need for the player to move from a fixed standing or seating position. As a result, there was no need to introduce players to the 'Guardian Player Protection System' designed by Oculus for player safety.

Participants gave informed consent by clicking the link and submitting the survey. They were informed that they were free to stop playing and leave the study at any point, should they wish, and were given details on data collection and processing. Once the link was clicked, they were taken to the next webpage. This contained the A-frame canvas with the game, a description detailing controls for the game, instructions for how the game is played, including the in-game challenge, a link to the corresponding survey to be completed after playing, and lastly were then provided with the link to the second version they would be due to play next. The players were instructed on how to start the game (enter full immersive VR mode) and how to play. They were free to play for as long as they wanted, using as many attempts as they wished (by refreshing the page) and were encouraged to play for around five minutes. Each version of the game required the player to fill out a linked survey hosted on Google Forms, with the same survey questions asked for each version of the game. The website for the study was hosted on Github pages.

In order to recruit participants with the most relevant background or interest, the study was promoted the study in specific targeted locations, such as the OculusQuest and VRGaming subreddits, the Facebook Quest gaming forum, The Creative Computing and Games Programme mailing lists, and direct messaging former music alumni and colleagues known to have an interest in gaming.

The first section of the survey was similar to the preliminary survey's pre-survey questions. It contained questions relating to demographic characteristics, experience with music performance and electronic music software, and experience playing video games. This can be found on page 243 of the appendix. The second section of the survey contained seven questions written by this researcher. It contained specific questions about the game they had just played, such as 'How responsive did the sound feel?' answered on a five-point Likert scale. The third section of the survey contained the established (and validated) Immersion question (IEQ) (Jennett et al., 2008). This is because the IEQ contains the most relevant questions for measuring immersion in a gaming experience as previously described in the Preliminary Study chapter on page 76. As the 31 questions were used in the preliminary study survey, this also allows a comparison between the two different prototypes.

Each survey group was exported from Google Forms in CSV format. They were combined in in a single Excel sheet with an additional label variable added to identify which group it was (either A or B). The questions relating to demographic characteristics and prior musical and gaming experience were separated from the survey data in Likert format. Additionally, the multiple-choice questions from the survey were also filtered out; 'If you had to pick one or the other, do you feel like it was primarily a musical experience or gaming experience?', 'Where did you place your attention?' and 'Did you reach the end of the platform?' The data from these questions and the questions pertaining to demographic characteristics will be addressed outside of the Paired-Sample testing separately. Lastly, the 10-point Likert scale question used as a baseline on the IEQ questionnaire for that participant was also removed from the analysis. The remaining data was all on a five-point Likert scale. In a similar fashion to the Preliminary Study, these remaining questions were split into three different sections. One section comprised only the questions written by the researcher, one comprised only the 31 IEQ questions, and the third with both sets of questions combined. It is also worth noting for questions 6, 8, 9, 10, 18 and 20 of the IEQ, a negative scale was used, and the results for these were reversed before being exported (so a '1' = '5' and vice versa). After the data was cleaned, it was imported into SPSS.

For each section of the survey data, Cronbach's alpha was calculated to measure the internal reliability. The section comprised of the four questions by this researcher had a Cronbach's alpha of 0.536, which could be considered moderately reliable (Hinton et al., 2004), given there are only a small number of questions. However, by comparing the Cronbach's alpha value against each individual question's value, it was found that the question 'How Challenging was the game?' did not correlate strongly with the scale overall; by removing this question, we can improve the internal consistency to 0.707, which is an acceptable level. A potential reason this question was found to be inconsistent could be the basis that everyone reported the game as equally hard (mean 3.33), regardless of how they answered the other questions. The version comprised of the 31 questions of the IEQ had a Cronbach's alpha of 0.893, showing great consistency. By comparing Cronbach's alpha value against each individual question's value, the following questions could be removed as they show the alpha value would increase if deleted (across each measure taken). As removing redundant questions found to measure the same behaviour reduces the chance of random error (Cronbach, L., 1951) (Koçak et al., 2014).

Q12: To what extent did you feel consciously aware of being in the real world whilst playing?
Q14: To what extent were you aware of yourself in your surroundings?
Q15: To what extent did you notice events taking place around you?
Q19: To what extent did you feel that the game was something you were experiencing rather than something you were just doing?
Q23: To what extent did you find the game challenging?
Q26: To what extent did you find the game easy?
Q28: How well do you think you performed in the game?
Q34: To what extent did you enjoy the graphics and the imagery?

After removing these questions, Cronbach's alpha for the remaining 23 questions was 0.924, which shows excellent internal consistency. However, as discussed in the Preliminary Study, it is worth considering that Cronbach's alpha could be high as a result of redundant questions (Tavakol & Dennick, 2011). However, these questions contain three of the most redundant questions for this research, given that the participants are wearing a headset and playing a VR game. Therefore, the validity of our internal consistency is accurate. Participants are unlikely to differ in reporting how aware they are in the real world, a change in awareness of their surroundings, and notice what events are taking place around them as their visual and auditory sensory environment is considerably closed when using the headset. The other inconsistent questions may also have a simple explanation. 'To what extent did you feel that the game was something you were experiencing, rather than something you were just doing?' could be considered difficult to interpret by participants, as it is potentially difficult to place on a scale. The three questions removed relating to the game challenge and performance potentially highlight the difficulty of this challenge. Lastly, given that the prototype was devised to test gaming mechanics devised for sonic expression, it is not surprising the minimal graphic style had minimal impact on the reported playing experience.

The Cronbach's alpha for the combined researcher's questions and IEQ is 0.901, showing excellent consistency. This is optimised to 0.926 when the inconsistent questions discussed above are removed, as described in the Methodology Chapter on page 60. The results from these 26 questions were then totalled and divided by the number of questions (26). Totals for each separate part of the survey were also generated. Section one (consisting of my questions) is the first four questions, part two being the remaining 23 questions from the IEQ. The results of the complete 31-question IEQ will also be discussed relative to the questionnaire's model; this will be examined separately.

Data Analysis: Combined Data

To determine the most appropriate test to use for comparing means, the distribution of the data must be checked for normality (Löfgren n.d.) (Field et al., 2013). Initially, this is performed by checking the histogram and Q Plot for the two separate groups (A and B versions of the game) by eye. The two histograms and Q Plots can be seen in the figures below.

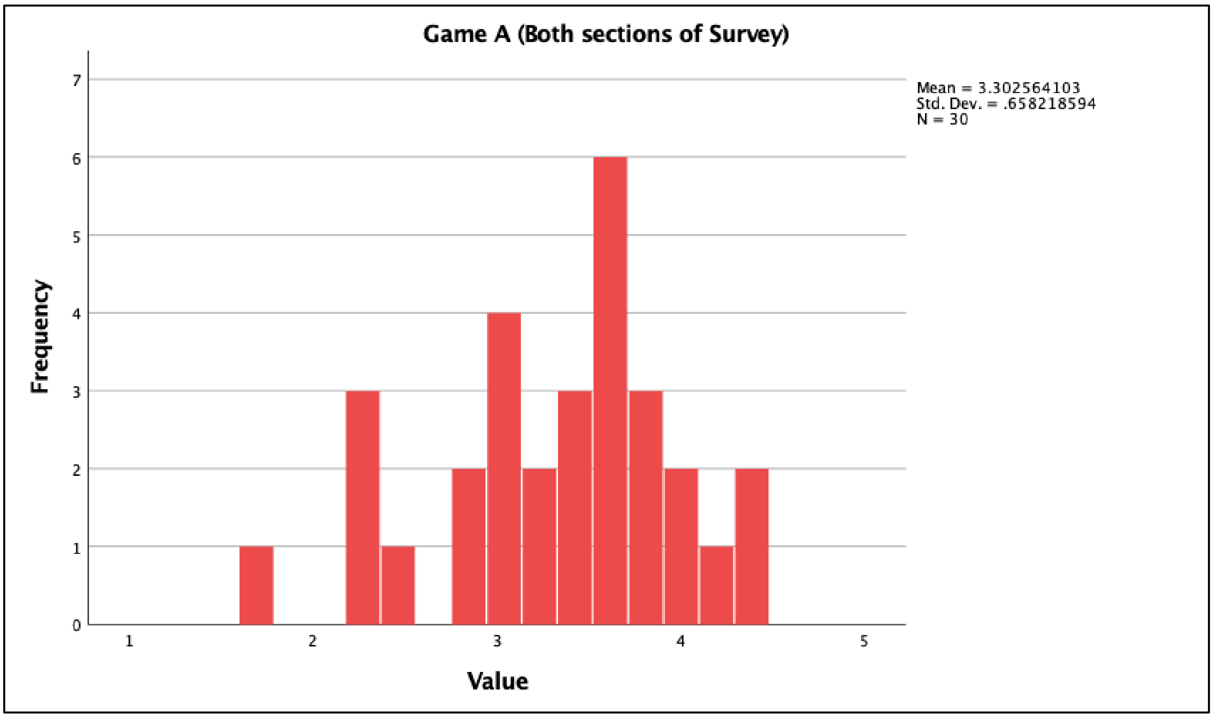


Figure 24: Study Two Game A Histogram Results for Complete Survey

We can see in the distribution of the histogram for Game A (the version where players had more agency in sample placement) that the distribution is skewed above a Mean value of 3.3. However, it does resemble a normal distribution. We can check the results further by calculating the Z-values for Skewness (S =-0.537, SE=0.427) and Kurtosis (S=0.272, SE=0.833), both were found to be within the acceptable range of -1.96 and 1.96 for normality (Löfgren n.d.). This is further supported by the group of values shown in the Q-Plot Figure 14 below.

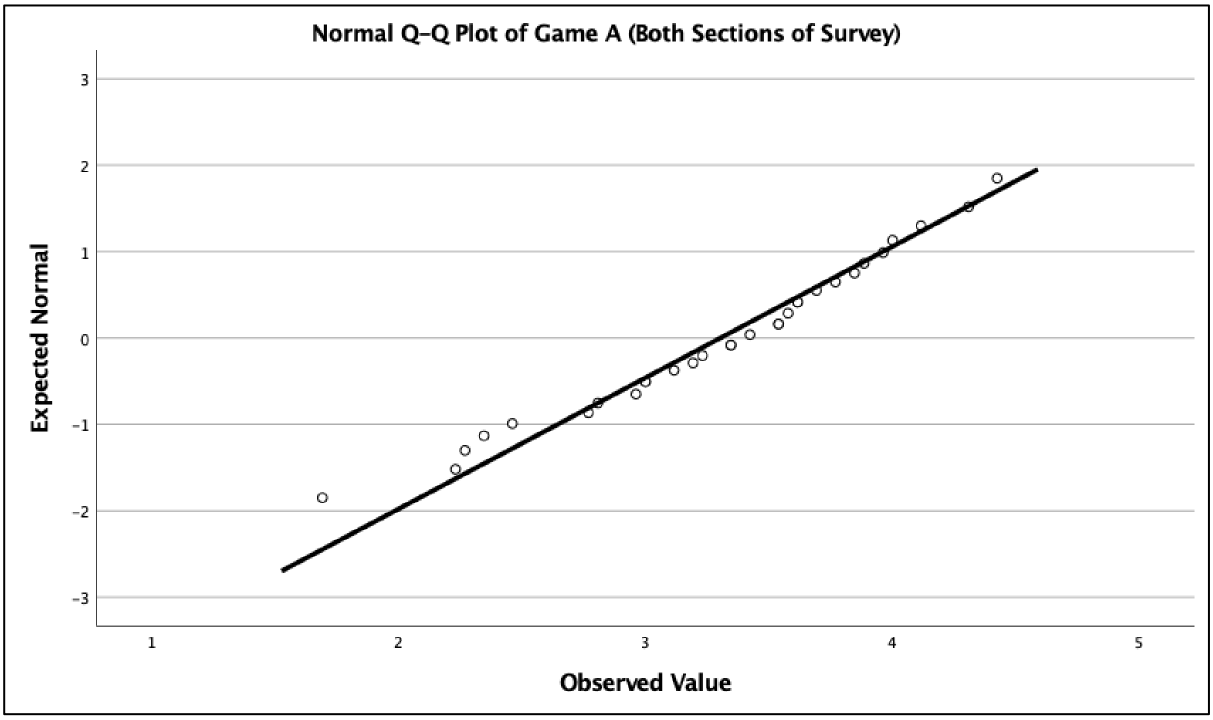


Figure 25: Study Two Game A Q-Plot for Complete Survey

There is a similar distribution in the histogram for Game B, where players had less control over note placement, in Figure 15 below, albeit distributed around a lower Mean of 2.93. These results were also checked by calculating Z-values for Skewness (S = 0.465, SE = 0.427) and Kurtosis (S = 0.272, SE = 0.833); both were found to be within the acceptable range (-1.96 to 1.96) for normality (Löfgren n.d.).

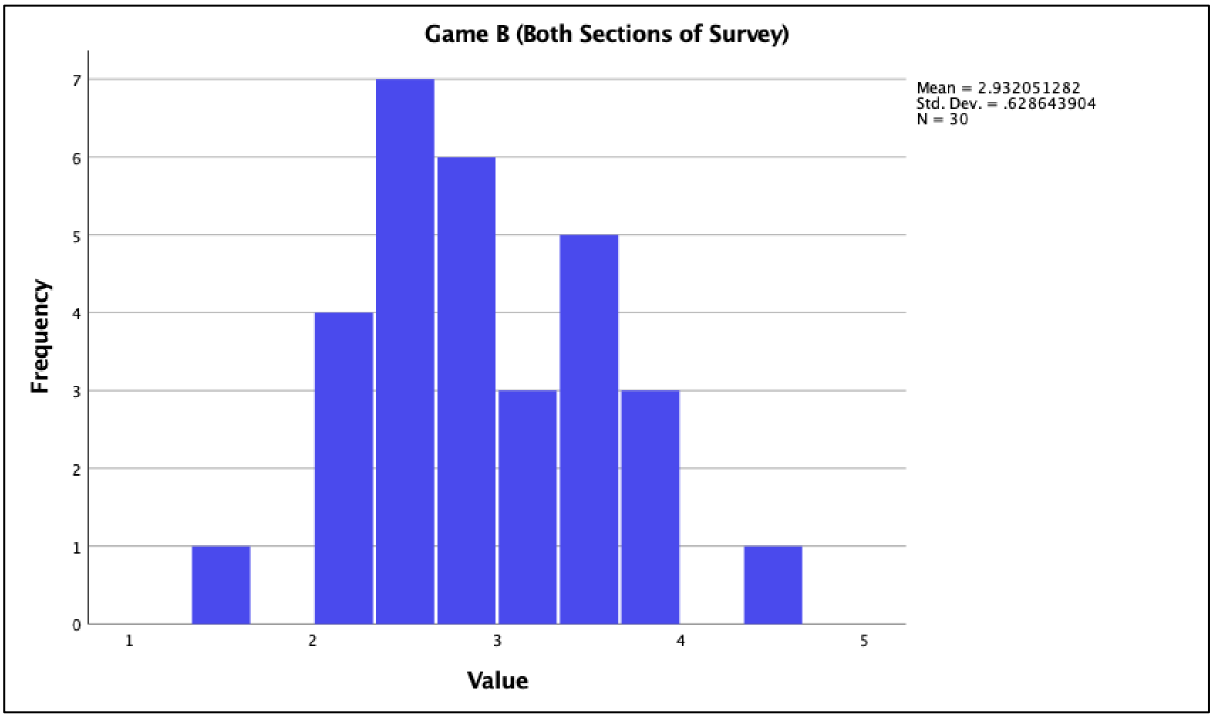


Figure 26: Study Two Game B Histogram Results for Complete Survey

This is further supported by the Q-Plot for Game B in Figure 16 below.

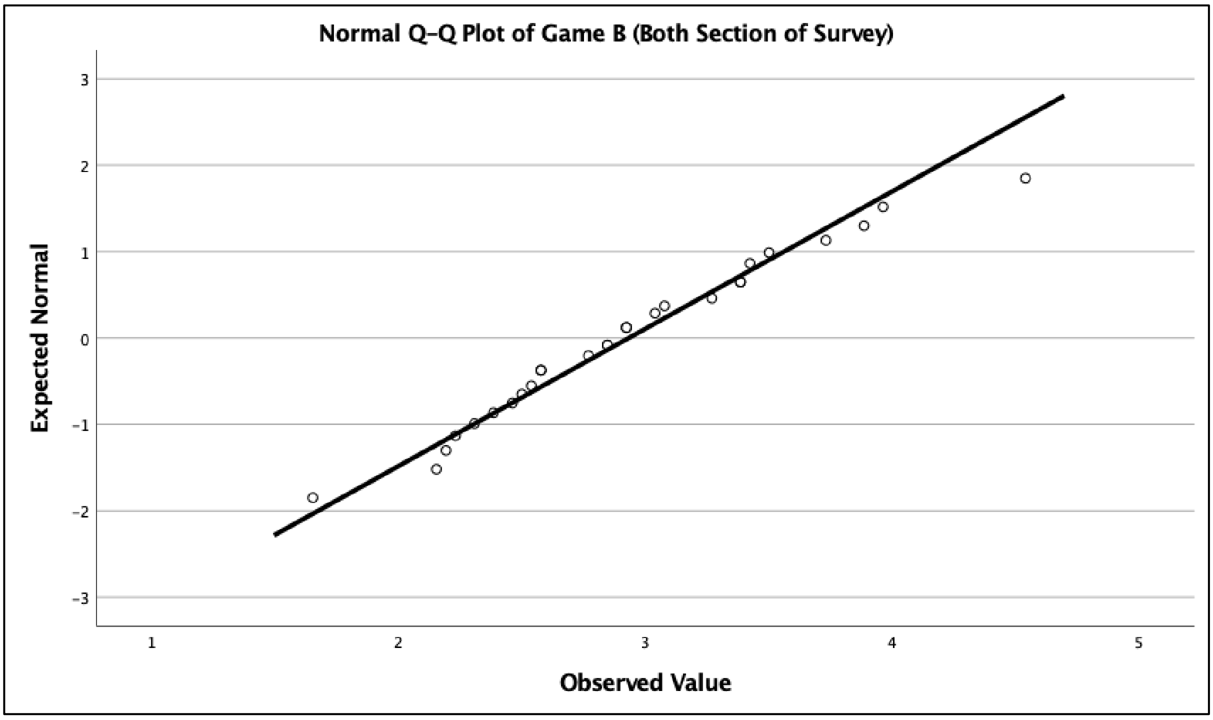


Figure 27: Study Two Game B Q-Plot for Complete Survey

As the sample size is small, a Shapiro-Wilk test ($p > 0.05$) (Shapiro & Wilk, 1965) of normality was performed, which also showed the null hypothesis to retain the 0.05 level of significance, therefore assuming normality for the dataset (Game A = 0.597, Game B = 0.753). As both groups contain the same participants, use the same surveys, and have normally distributed results, it has met the necessary assumptions (Field et al., 2013) for the Dependent-means/Paired-Samples T-test to be the most appropriate method of comparing means between the playing experience of both versions.

Table 13: Paired Sample Statistics for 2nd Study Survey (Combined)

Paired Sample Statistics					
		Mean	**N**	**Std. Deviation**	**Std. Error Mean**
Pair	GameABothParts	3.30256410250	30	.658218593876	.120173723878
	GameBBothParts	2.93205128210	30	.628643903856	.114774148927

The label 'GameABothParts' means the pair contains the results from both the section of the survey written by me and the section containing the IEQ for the Game A version, whereas the label 'GameBBothParts' contains the results from both corresponding sections for the Game B version.

Table 14: Paired Samples Correlations for 2nd Study Survey (Combined)

Paired Samples Correlations				
		N	**Correlation**	**Sig.**
Pair	GameABothParts & GameBBothParts	30	.618	.000

Table 15: Paired Samples Test for 2nd Study Survey (Combined)

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	**Std. Deviation**	**Std. Error Mean**	**Lower**	**Upper**	**t**	**df**	**Sig. (2-tailed)**
Pair	GameABothParts - GameBBothParts	.370512820400	.563050119507	.102798417153	.160266450474	.580759190326	3.604	29	.001

Cohen's d was used to calculate the effect size (Wiseheart n.d) (Cohen, 1988). This means that, on average, participants had a significantly greater player experience and immersive experience in the Game A version, where they had more expressivity, compared to the less expressive Game B version (M = 0.371, SE = 0.563, $t(29) = 3.60$, $p < 0.05$, with an effect size d = 0.658). Given that the full survey contains questions relating to both the playing experience and the player's sense of immersion, it would be beneficial to see how this compares to each individual section of the survey and relates to each group.

Data Analysis: Section One

The distribution for the three questions written by the researcher can be found in Figures 17, 18, 19, and 20 below. Normality was checked as previously described.

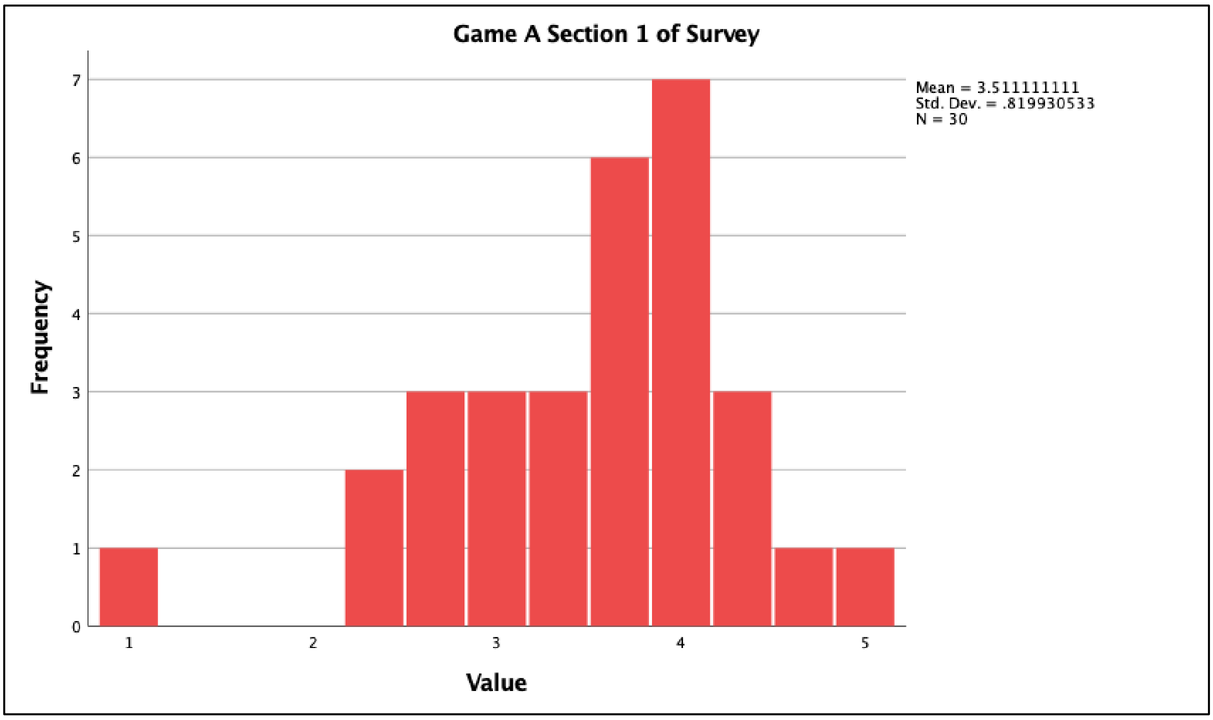


Figure 28: Study Two Game A Histogram Results for Section One of the Survey

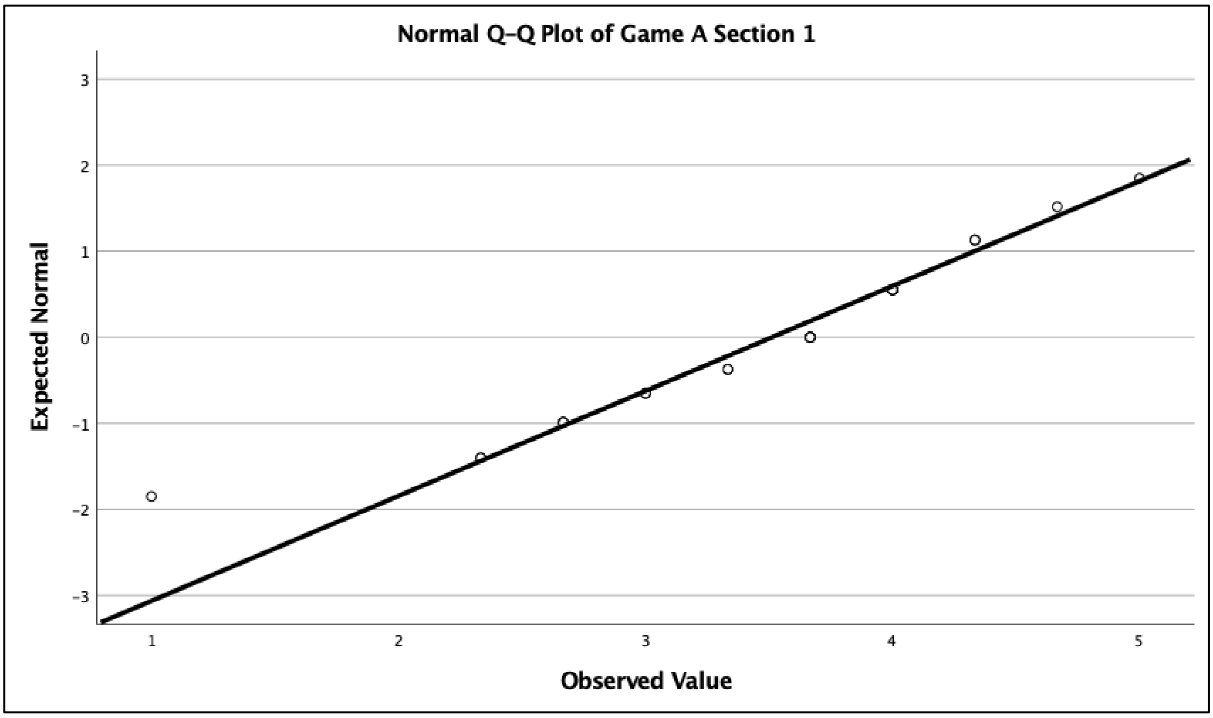


Figure 29: Study Two Game A Q-Plot for Section One of the Survey

The distribution for section one of Game A, as approximated by the eye, looks similarly normal as those of the overall results. However, they are slightly more concentrated around a slightly higher mean of 3.511 compared to the combined Game A results (M = 3.303). There appears to be a single value, potentially an outlier.

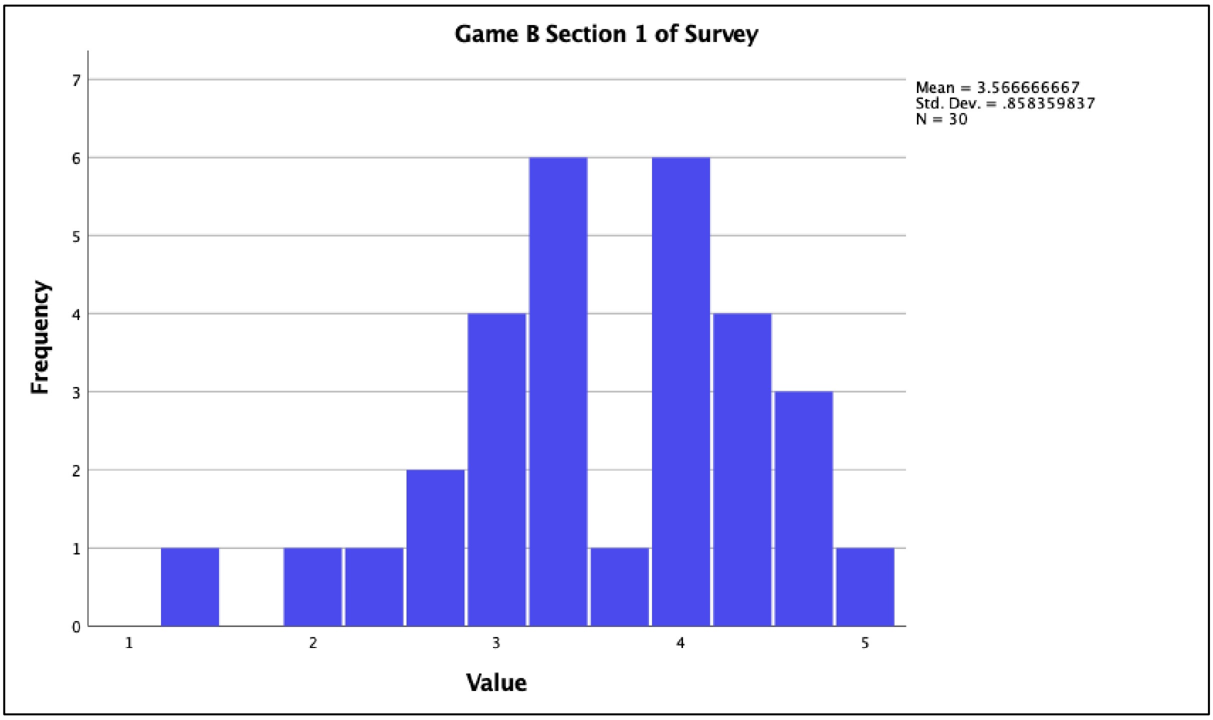


Figure 30: Study Two Game B Histogram Results for Section One of the Survey

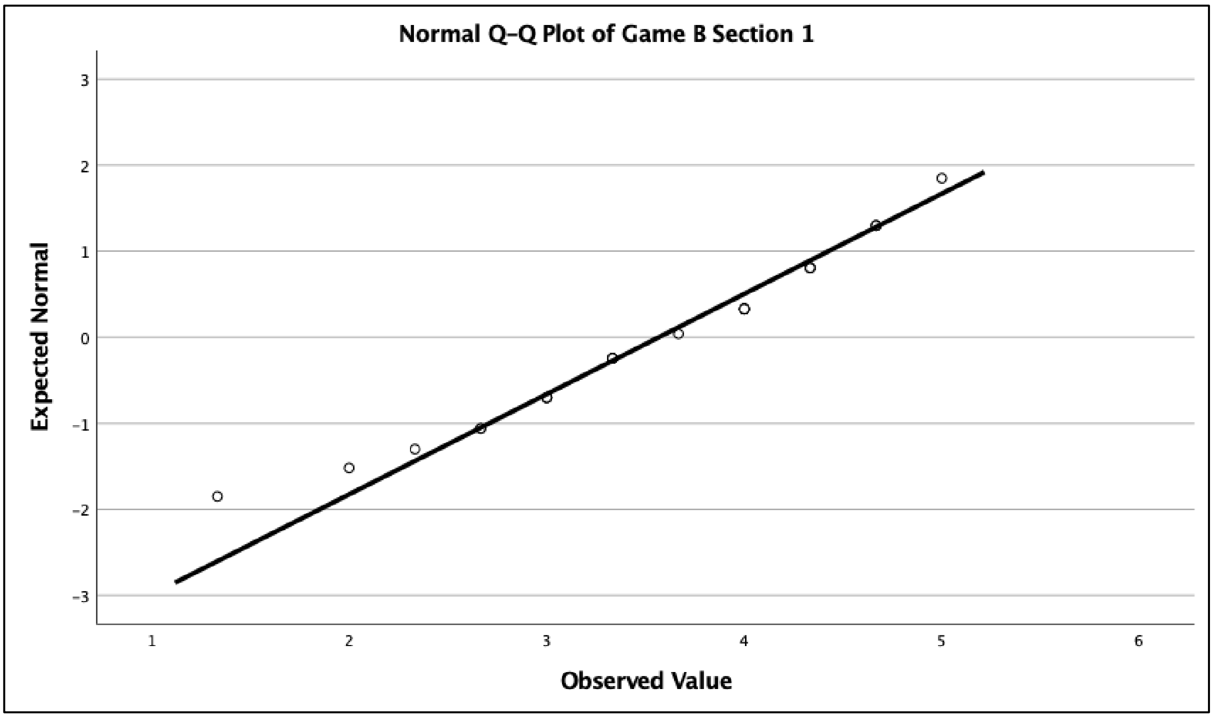


Figure 31: Study Two Game B Q-Plot for Section One of the Survey

The distribution for the Game B section one results, as approximated by the eye, also appears to be similarly normally distributed in reference to the combined results. However, the values are concentrated in narrow bands of values with a comparatively higher mean of 3.567 than that of the combined Game B results (M = 2.932). After further checking the z-values for Skewness for both Game A (S = -0.937, SE = 0.427) and Kurtosis (S = 1.713, SE = 0.833) and Game B Skewness (S = -0.578, SE = 0.427) and Kurtosis (S = 0.192, SE = 0.833) this was confirmed and found to be in an acceptable range of normality (Löfgren n.d.). Using the Shapiro-Wilk test again, the null hypothesis retains the 0.05 level of significance, therefore assuming normality (Game A p = 0.076, Game B p = 0.267).

A Paired-Samples T-test was used to compare means:

Table 16: Paired Sample Statistics for 2nd Study Survey (Section One)

Paired Sample Statistics					
		Mean	**N**	**Std. Deviation**	**Std. Error Mean**
Pair	GameAPart1	3.51111111113	30	.819930533144	.149698149530
	GameBPart1	3.56666666660	30	.858359836708	.156714348340

The label 'GameAPart1' means the pair contains the results for only the first section of the survey for the version of the game that allowed the players to choose their sample and placement, whereas the label 'GameBPart1' contains the first section of the survey for the version where players were not given a choice over sample and placement.

Table 17: Paired Samples Correlations for 2nd Study Survey (Section One)

Paired Samples Correlations				
		N	**Correlation**	**Sig.**
Pair	GameAPart1 & GameBPart1	30	.522	.003

Table 18: Paired Samples Test for 2nd Study Survey (Section One)

Paired Samples Test										
		Paired Differences								
					95% Confidence Interval of the Difference					
		Mean	**Std. Deviation**	**Std. Error Mean**	**Lower**	**Upper**	**t**	**df**	**Sig. (2-tailed)**
Pair	GameAPart1 - GameBPart1	-.055555555467	.821564405036	.149996452360	-.362332746049	.251221635116	-.370	29	.714

The results show a statistical significance result of 0.003, meaning the difference is not due to chance. The effect size was calculated using Cohen's d (Wiseheart n.d) (Cohen, 1988). Consequently, on average, participants reported a slightly more intuitive, responsive, and controllable experience in Game B, where the players had limited agency over their musical interactions. It is not surprising that players reported the controls felt slightly more intuitive than in Game A as the demand on the player's ability is lower ($M = -0.056$, $SE = 0.822$, $t(29) = -0.370$, $p < 0.05$, effect size $d = 0.822$). It is also not surprising they reported it as a more intuitive gaming experience, given that it requires less to plan and execute their input.

The results of the multiple-choice questions from the first section are also useful to inform our understanding of the player experience. 30% of players reported that Game A felt like a 'musical experience' over a 'gaming experience' compared to only 10% of Game B. Players also reported placing more attention on arranging sound (57.67%) over attempting to reach the end platform (33.33%) in Game A, with 10% unsure where they placed their attention. Comparing this to Game B, where attention was on arranging sound was only 6.66%, attempting to reach the end platform was 76.67% with 16.67% unsure, we can start to infer where the experience was different. Another interesting result from section one of the survey is players reported that they were more successful reaching the end platform in Game B (66.67% reported they reached the end) compared to fewer players (46.67%) reporting they reached the end of the level in Game A. This is also supported by the difference in 'How challenging' the players found the game: Game A was found to be less challenging (M = 3.77, SD = 1.073) compared to Game B (M = 2.9, SD = 1.242). Evaluating these questions and answers in relation to demographics may also prove useful in furthering understanding of player experience. This will be done after analysis of section two of the survey.

Data Analysis: Section Two

The distribution for the remaining 23 questions relating to player immersion was checked for normality as previously described. The analysis is below:

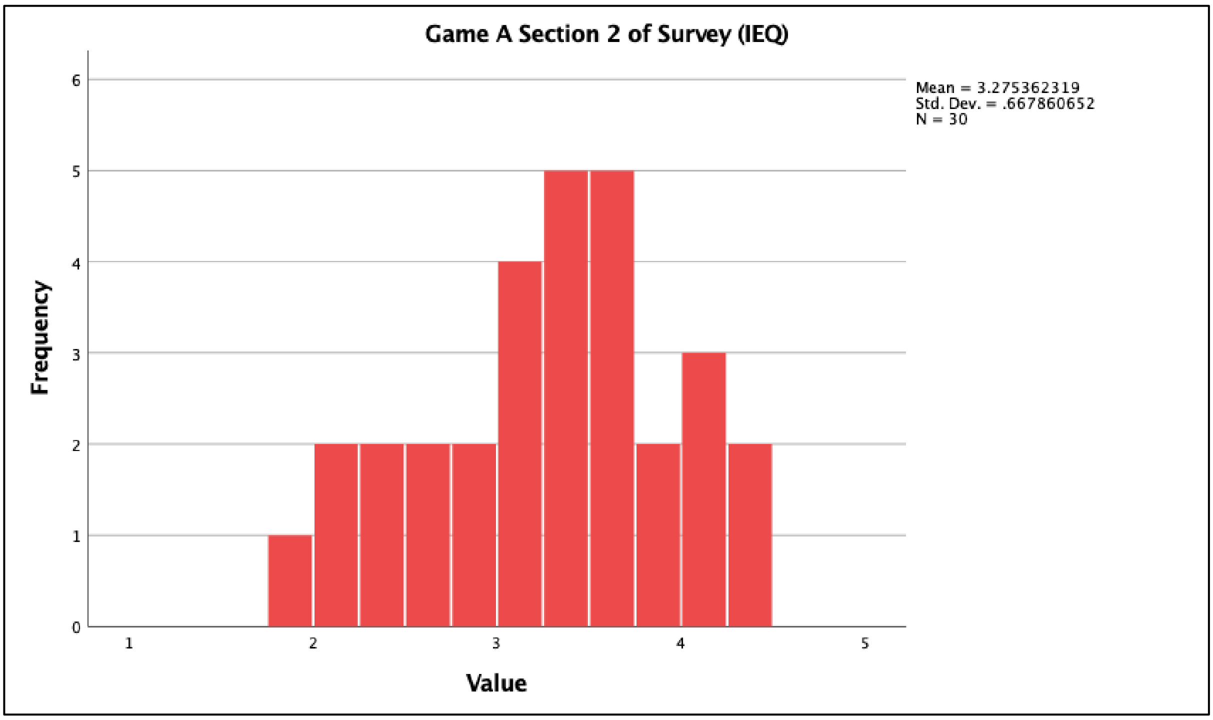


Figure 32: Study Two Game A Histogram Results for Section Two of the Survey

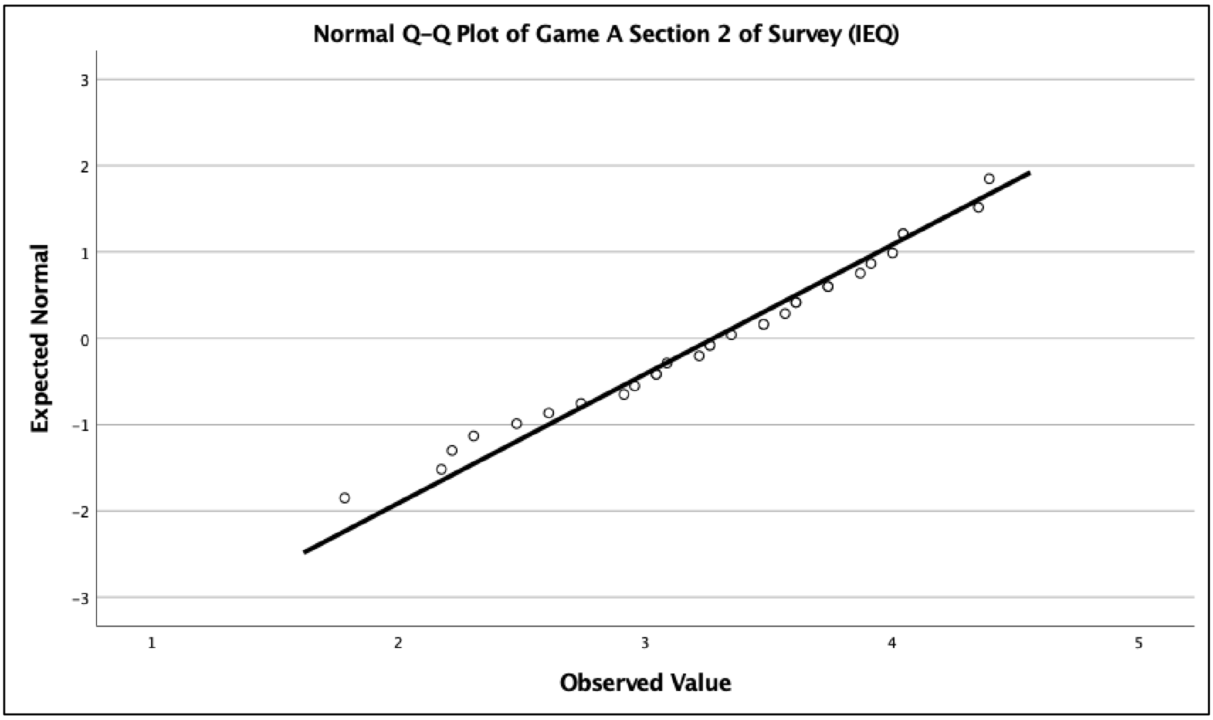


Figure 33: Study Two Game A Q-Plot for Section Two of the Survey

The distribution of the Game A Section Two results, as approximated by the eye, looks more normal and less skewed compared to the equivalent Section One and is similar in distribution to the combined results. After checking the z-values for Skewness (S = -0.389, SE = 0.427) and Kurtosis (S = -0.446, SE = 0.833), this was confirmed and found to be within an acceptable range for normality (Löfgren n.d.).

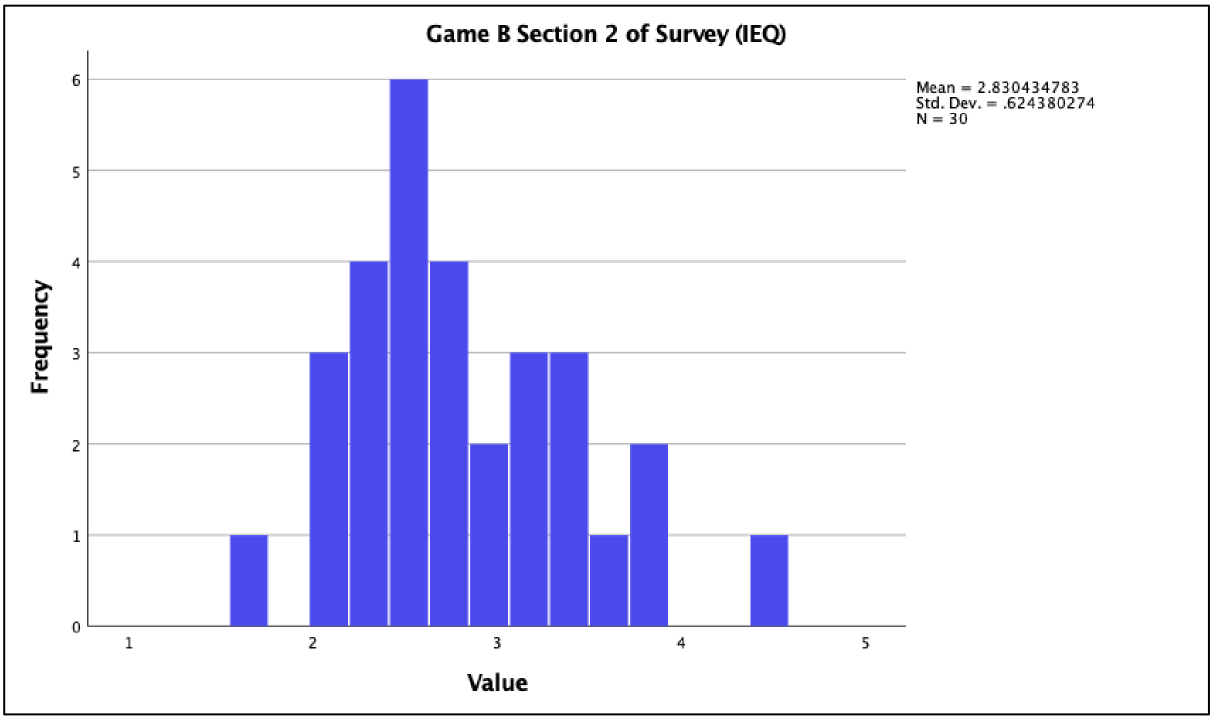


Figure 34: Study Two Game B Histogram Results for Section Two of the Survey

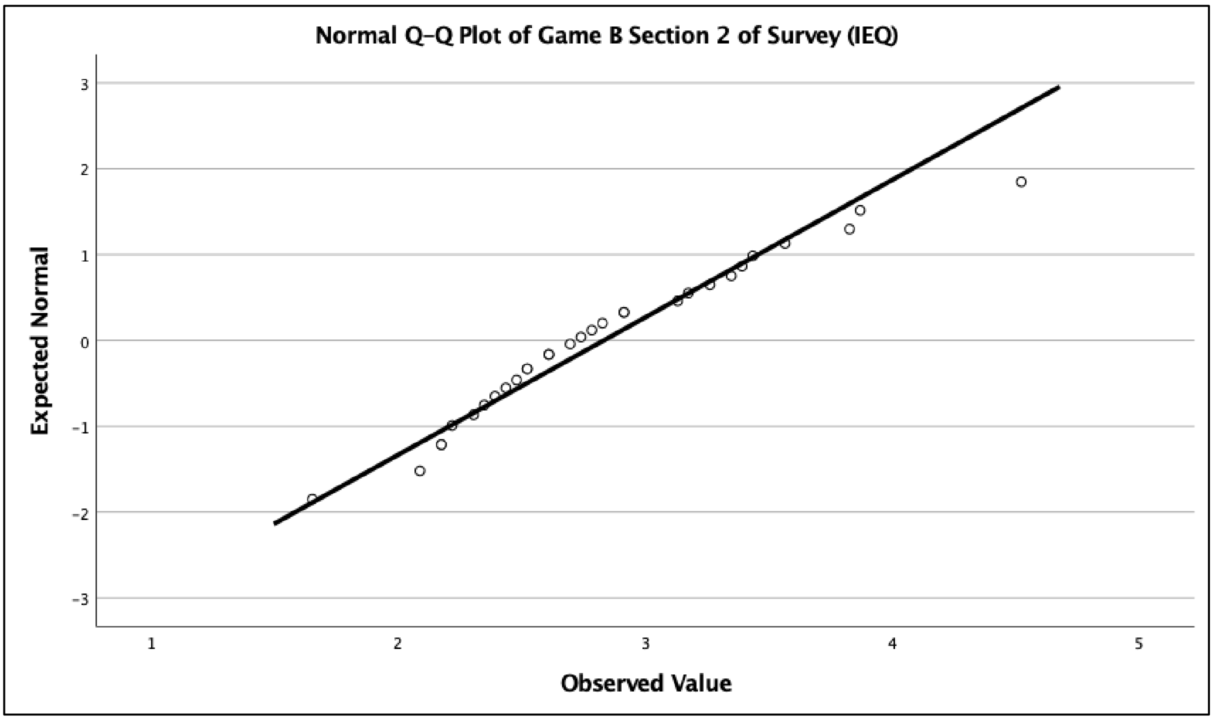


Figure 35: Study Two Game B Q-Plot for Section Two of the Survey

The distribution of the Game B Section Two results, as approximated by the eye, looks similarly slightly skewed but normally distributed when compared to the combined results. This is confirmed after checking the z-values for Skewness (S = 0.706, SE = 0.427) and Kurtosis (S = 0.559, SE = 0.833) and found within the acceptable range for normality (Löfgren n.d.).

The Shapiro-Wilk test also showed the null hypothesis to retain the 0.05 level of significance, therefore assuming normality (Game A p = 0.672, Game B p = 0.369).

A Paired-Samples T-test was once again used to compare means.

Table 19: Paired Sample Statistics for 2nd Study Survey (Section Two)

Paired Sample Statistics						
		Mean	N	Std. Deviation	Std. Error Mean
Pair	GameAPart2IEQ	3.27536231883	30	.667860652132	.121934114814
	GameBPart2IEQ	2.83043478257	30	.624380274195	.113995720213

The label 'GameAPart2IEQ' means the pair contains the results for only section two of the survey for Game A, the version of the game where players had control over sample choice and placement, whereas the label 'GameBPart2IEQ' contains the results only for section two of the survey for Game B, the version of the game where players had no control over sample selection and placement.

Table 20: Paired Samples Correlations for 2nd Study Survey (Section Two)

Paired Samples Correlations				
		N	Correlation	Sig.
Pair	GameAPart2IEQ & GameBPart2IEQ	30	.594	.001

Table 21: Paired Samples Test for 2nd Study Survey (Section Two)

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair	GameAPart2IEQ – GameBPart2IEQ	.444927536267	.583510159004	.106533892207	.227041262034	.662813810499	4.176	29	.000

The results show a statistically significant result of 0.001, meaning the difference is not due to chance. The effect size was calculated using Cohen's d (Wiseheart n.d) (Cohen, 1988). On average, participants had a significantly greater playing and immersive experience when playing Game A, the version with more agency over sample selection and placement, compared to Game B, the version with no control ($M = 0.445$, $SE = 0.584$, $t(29) = 4.176$, $p < 0.05$, with an effect size $d = 0.584$). The second section of the survey is best understood as a whole. However, there are a few suitable questions that can be observed to give further insight into the playing experience to accompany the first section. Question 19: 'To what extent did you feel motivated while playing?' In Game A ($M = 3.5$, $SD = 0.938$), players reported they were more motivated than in Game B ($M = 2.767$, $SD = 1.073$). Question 25: 'How much did you want to win the game?' The desire to win both versions was very similar for Game A ($M = 3.1$, $SD = 1.2690$) compared to Game B ($M = 3.133$, $SD = 1.195$). Question 29: 'How much would you say you enjoyed playing the game?' More participants were positive In Game A ($M = 3.367$. $SD = 1.189$) compared to Game B ($M = 2.767$, $SD = 0.858$). Question 31: 'Would you like to play the game again?' Participants were in favour of Game A ($M = 3.4$, $SD = 1.248$) than Game B ($M = 2.5$, $SD = 0.9738$).

Data Analysis: IEQ Scoring

The results from the original 32 IEQ questions (31 five-point Likert and one 10-point baseline Likert) are documented herein. The IEQ provides scores for five immersion factors: 'Cognitive Involvement', 'Emotional Involvement', 'Real-World Dissociation', 'Control', and 'Challenge' (Jennett et al., 2008). As previously described in the Preliminary Study Chapter on page 90, the totals for each factor have been normalised to the one-to-five scale of the Likert scale. This allows for direct comparison between factors as they are appropriately scaled. Individual participant IEQ scores for both Game A and Game B surveys were checked against their baseline, with all being found comparable. Additionally, this can be communicated in the form of the total IEQ score for Game A ($M = 3.29$, $SE = 0.538$) compared to the total baselines for Game A ($M = 3.15$, $SE = 0.957$) and total score for Game B ($M = 2.968$, $SE = 0.503$) compared to totalled baseline for Game B ($M = 2.85$, $SE=0948$).

The following table contains the breakdown per immersion factor for Game A, with the baseline mean normalised to the five-point Likert scale.

Table 22: IEQ Scoring Immersion Factors for Game A (Study Two)

IEQ Scoring Immersion Factors							
	Cognitive involvement	**Emotional Involvement**	**Real World Dissociation**	**Control**	**Challenge**	**Overall**	**Baseline**
Mean	3.533	2.622	3.629	3.386	3.333	3.329	3.15
Std. Deviation	0.702	0.780	0.707	0.748	0.6100	0.538	0.957

The following table below contains the breakdown per immersion factor for Game B with the baseline normalised as above.

Table 23: IEQ Scoring Immersion Factors for Game B (Study Two)

IEQ Scoring Immersion Factors							
	Cognitive involvement	**Emotional Involvement**	**Real World Dissociation**	**Control**	**Challenge**	**Overall**	**Baseline**
Mean	3.281	2.1	3.467	3.02	2.625	2.968	2.85
Std. Deviation	0.680	0.618	0.574	0.659	2.625	0.503	0.948

We can see in the tables above, that while participant baselines were under-reported compared to their IEQ score in both versions of the game, their scores are still close. This is not surprising as we can expect the outcome of 31 answers to be more detailed than a single baseline question. All five factors reported an increase in immersion in Game A. This is a larger increase in immersion than found by Thompson in the study of effect of touch-screen size on game immersion (iPad (M = 3.761) compared to an iPod Touch (M = 3.3)) (Thompson et al., 2021) and found in Cox's study of the role of challenge in the gaming experience; where immersion reportedly decreased for low expertise players when given a high effort challenge (normalised low expertise player M = 3.99, high expertise player M = 3.693) (Cox et al., 2012). We can see from both versions of the game that 'Real World Dissociation' scored highest, although it differed the least between the two versions of the game. This is not surprising given the enclosed nature of the VR headset, as discussed in the previous chapter containing the Preliminary Study on page 76, as both versions of the game were delivered on the same device. However, the other four categories offer some interesting results. The players had more agency regarding how they navigated and interacted with the game environment in Game A, requiring more control input and thought-out decision-making. One could argue that this, in turn, increases the 'challenge' component and necessary cognitive involvement in performing the task, reflected in the player's increased sense of immersion in the table above. The potential relationship between these varied factors and the rest of the results will be discussed in more detail in the evaluation below.

Demographics

The salient demographic questions asked were pertaining to gaming experience and musical experience. Of the 30 participants, 26 (86.7%) had played games for more than five years, with three (10%) having played for more than a year but less than five years. The remaining one participant (3.3%) had played for more than a week but less than a month. Despite this potential disadvantage, this participant scored well within the spread of the standard deviation. Only four participants usually played on an Oculus headset, with one participant having played on a PSVR. The results show this did not confer any notable benefit. The most common device was a console without motion controls, with 19 of the 30 (63.3%) usually playing that type of device. Only 14 out of 30 (47.7%) played with motion controls at some frequency, with nine of those stating a couple of times a year. The remaining 16 out of 30 (53.3%) claimed to never use motion controls. The 14 participants who had some experience with motion controls found the interaction more intuitive (M = 3.571, SD = 1.213) and responsive (M = 3.571, SD = 1.004) compared to the 16 participants who had no experience; intuitive interaction (M = 3.06, SD = 0.681), and responsive (M = 3.25, SD = 0.775). Interestingly, there was little difference between how in-control players felt, regardless of experience with motion controls.

18 out of 30 participants (60%) had experience playing a musical instrument. Of these 16, nine (50%) had played for more than seven years. Only two participants had played for less than a year. Those with seven years or more of musical instrument experience reported they found it less challenging (M = 3.111, SD = 1.101) compared to those with no musical

instrument experience (M = 4, SD = 0.954). They also reported feeling slightly more in control with experience (M = 4.333, SD = 0.843) compared to those without (M = 4.08, SD = 1.068). 50% of the participants had experience using music production software. 11 out of these 15 (73.33%) had experience of playing a musical instrument. There was no notable difference in how the 50% with music production experienced the game, although they reported they felt more in control (M = 4.133, SD = 1.125) compared to those without experience (M = 3.357, SD = 1.277). Interestingly, those with music production experience were proportionally more successful in reaching the end platform in both versions of the game. In Game A, eight of the 14 players that reached the end platform had music production software experience. In Game B, 10 out of the 20 players that reached the end platform had music production software experience. This is interesting, as it is potentially possible that having some practical experience or a grounded conceptual understanding of how a step sequencer worked benefitted the successful completion of the game.

Evaluation

The results between the two distinct groups across all sections of the survey indicate the playing experience and the player's sense of immersion changed significantly, as shown in Table 21 and Table 22 on page 122. In the first section of the survey regarding playing experience, there was only a small amount of difference in how intuitive the interaction was or how responsive the sound was between each version of the Game in favour of Game B (Game A M = 3.511, Game B M = 3.566). However, one would anticipate this situation, given there is a higher demand for interaction and spatial reasoning in Game A than in Game B. The results for questions relating to 'cognitive involvement' in section two reflect players had an increased level in Game A (M = 3.533) when compared to Game B (3.281), which would support this reasoning. Additionally, there is a broader range of potential sonic events, which could account for the difference in how responsive the sound felt (Game A M = 3.4, SD = 0.894, Game B M = 3.833 SD = 0.949). Sound is not triggered automatically upon collection in Game A; it requires the player to place the note. Once placed, it will only trigger when the playhead reaches that point in the loop, as opposed to the immediacy of the trigger upon collision/collection in Game B. This change in gaming mechanics could be responsible for the apparent change in responsiveness.

Another factor could be the difference in the computational overhead required between the two different versions of the Game. To achieve the desired effect programmatically, Game A required a larger codebase with more calculations performed at runtime in the browser on the headset. Subsequently, it is subject to the limitations and hierarchical structure of how the browser chooses to organise and prioritise the device’s hardware resources. If the user were to be impacted by the allocation of memory (and garbage collection), it would more likely be in Game A. Attempts were made to mitigate this by optimising the codebase, minimising the amount of data and assets being allocated and destroyed at runtime. Where possible, pools of assets were used, calls were minimised, and Document Object Model (DOM) access was reduced.

Additionally, there are questions in the second section of the survey pertaining to player experience that highlight the playing experience may still be greater in Game A. Players reported a greater motivation to continue to play (Game A M = 3.5, Game B M = 2.767), a desire to play again (Game A M = 3.4, Game B M = 2.5), and a higher sense of enjoyment in Game A (M = 3.367), compared to Game B (M = 2.767). The difference between the two versions of the games across these three factors is substantial. One can also state confidently that these are significant factors for playing experience. If players are not motivated to continue playing, they will stop, and the playing experience will cease. If players do not have a desire to play again after finishing, it speaks to the *replayability* (and subsequent value) of the playing experience. If players do not find the experience enjoyable, will they continue to play, recommend, or engage with the game positively at all? Interestingly, the player’s sense of enjoyment did not come at the expense of finding the game easy. Players found Game A noticeably more challenging (M = 3.77) compared to Game B (M = 2.9). We could claim the implicit self-imposed challenge of musical expression (i.e., the demand of making musical decisions) is a reasonable factor for this increase, along with the increased physical and cognitive demand of matching the intended outcome with the appropriate input and performance. Perhaps unsurprisingly (given that the players reported Game B to be less challenging than Game A), more players were able to reach the finish in Game B, with 66.67%

(20 out of 30 players) successful compared to 46.67% (14 out of 30 players) capable in Game A.

Despite finding it harder to complete the game, the level of enjoyment was higher for Game A. Subsequently, we could claim that when given a task with the increased agency of sonic expression, players will value this implicit self-imposed challenge of sonic expression more than the outcome of an explicit in-game challenge. Overall, we can argue that giving players more expressivity in their sonic arrangement coupled with more agency in how they can solve the gaming challenge, leads to a significantly greater experience for the player. There is no trade-off in terms of the player's desire to win (Game A M = 3.1, Game B M = 3.133), although it does come with a slightly harder challenge to the player. As mentioned in the data analysis of the demographic above, this could be mitigated with some prior experience in electronic music arrangement.

The same argument can be made for the impact of greater agency of sonic expression leading to a significant increase in the player's sense of immersion (Game A section two M = 3.28, Game B section two M = 2.830). The results of the second section and the complete IEQ make a compelling argument that player immersion can be increased by expanding the player's ability to express themselves with sound or music. The results of the full 31-question IEQ show players felt more immersed across all five immersive factors. For 'Cognitive Involvement' (Game A M = 3.533, Game B M = 3.281), we can infer this change is due to the additional cognitive demand of making an additional decision to navigate the game environment, coupled with the fact it is a musical decision with an additional dimensionality of spatial reasoning, along with the corresponding increase of interaction demand this action requires. We can further support this argument by looking at the difference between those with musical production software experience and those without (as we had an even split), as prior experience could benefit cognition. In Game A, those with music production experience reported an increase in Cognitive Involvement (M = 3.511) compared to Game B (M = 3.160). The same increase in Cognitive involvement from those without experience was also reported higher in Game A (M = 3.55) compared to Game B (M = 3.166). Interestingly, regardless of experience, the level of Cognitive Involvement remained similar between these subgroups.

Subsequently, we can argue that prior music production experience does not prove substantially beneficial in eliciting Cognitive Involvement in player immersion in our game.

In the IEQ factor 'Emotional Involvement' (Game A M = 2.622, Game B M = 2.1), we could infer that the increase is due to the increasing sonic expression of the player coupled with the affective nature of sound. We could argue that players had a stronger sense of emotional immersion as they were more responsible in the sonic experience. The difference between participants with prior music production experience and those without was minimal, but in both subgroups for each version of the game, those with no experience scored the sense of immersion slightly higher (Game A with experience M = 2.589, those without M = 2.655, Game B with experience M = 2.042, those without M = 2.167). This says something interesting about the relationship between comprehension, sensory experience, and affective sound, although it is out of the remit of this research. As discussed in the data analysis for the IEQ, the factor 'Real World Dissociation' is very similar across both versions of the Game (Game A M = 3.629, Game B M = 3.467). This is likely due to the nature of wearing a headset and obstructing physical surroundings. The difference between the two game versions could be explained by the impact of motivation, as discussed above in the evaluation of playing experience. If players are not as motivated to continue playing, they would likely be more inclined to remove their headsets or express this desire in their surroundings.

In the IEQ factor of 'Control', we see that players felt a noticeable increase in immersion when given more control over their interactions (Game A M = 3.386, Game B M = 3.02). This is interesting as it is contrary to what was reported in the playing experience. While players found the controls were slightly harder in Game A, impacting the playing experience negatively (albeit comparably not by much), it has an adverse effect on their sense of immersion for both versions of the game. Is there a relationship between perceived responsiveness and control? This could be explored within future studies. There was also a difference in results between subgroups of those with music production experience and those without. In Game A, those with experience (M = 3.44) reported a greater sense of immersion compared to those without (3.33). In Game B, those with experience (M = 2.975) reported less than those without (M = 3.07). In both instances, the difference is small. However, one could argue Game A is more demanding regarding control. Subsequently, it is more likely prior contextual experience of a step sequence may serve to be more useful in Game A than in Game B. Additionally, as Game B follows a more traditional music game paradigm of sequence matching with little player agency, those without experience would not be disadvantaged, with gaming experience being a more crucial factor.

In the final IEQ factor 'Challenge', players reported a substantial increase in immersion (Game A M = 3.333, Game B M = 2.625). In the subgroups of those with music production software experience and those without, in both instances of the game, players with no experience reported a higher sense of immersion for the challenge (Game A with music production experience M = 3.317, without M = 3.35. Game B with music production experience M = 2.578, without M = 2.679). Although the intra-difference is small in both subgroups, one could argue that this slight increase in the challenge factor comes from novelty in task engagement. This is supported by the results of section one of the survey, where players without music production experience found the playing experience more challenging. An increase in the sense of challenge could explain a higher engagement and engrossment in the task, as those with prior transferable experience may find the task less novel and subsequently engaging.

Conclusion

By giving players greater agency in sonic expression, one can increase the player's sense of immersion significantly. This also impacts the playing experience substantially in relation to player motivation and enjoyment. However, by increasing player agency in this domain, there is an increase in how challenging players find the experience. Although some relevant experience of musical concepts or electronic music production can be beneficial for the players, it was not overly beneficial for player enjoyment. This balance between agency, challenge, and enjoyment will be explored in greater detail in future studies. To what extent do these three factors impact each other? To what extent does the interplay between agency, sonic outcome, and challenge influence immersion? How limited is too limited with respect to expression and the resultant immersion? Answering these questions will provide greater detail in understanding the role sonic expression has on the playing experience and immersion.

Study Three: Music Game Design and the Impact of Non-Diegetic Synchronised Cues on the Playing Experience

Abstract

This chapter presents a study on synchronised sound design and musical arrangement in an experimental VR music game called 'Grail'. In this game, the players must shoot down all target blocks that move towards them and avoid being hit by them. Players are asked to play two versions of the game; in each version, the destruction of the target blocks triggers a different sound. The study explores the impact of sound design on the playing experience and the player's sense of immersion. This study will seek to ascertain the extent to which interactive sound design can encourage synchronised input and whether this can be utilised to make novel gaming experiences. It is hoped that the outcome of the study will establish whether player interaction can be utilised in sound design and whether increasing the significance of the sonic outcome derived from the player's interaction has an impact on their playing experience.

Introduction

This study was conducted to determine the importance of interactive sound-design decisions, specifically in the context of synchronised audio-visual game events. The study was a 'within group' study with two different versions of the game in an A/B testing format. A total of 42 participants played at least one version of the game, 26 of whom played both versions of the game. These 26 were used for the study. The 16 participants who only played one version of the game had their data discarded, as they would need to play both versions to provide meaningful data. This study focused on the exploration of different sound design choices and corresponding sonic output derived from player interaction, specifically in the context of the playing experience and the impact the sound design has on the player's sense of immersion. The aim of the game is for the player to shoot all the moving target blocks coming towards the player and reach the end of the piece of music, ideally without the blocks colliding with the player. The outcome of this study will improve our understanding of what impacts a player's sense of immersion. This includes whether different sound design choices, when used with a synchronised gaming challenge, can have an impact on the player's experience. It will also improve our understanding of the role of diegetic and non-diegetic sound design and to what extent players should be responsible for triggering such sounds with accurate timing.

Methodology

A 3D rhythm game was developed to provide a typical gaming challenge. The game contained features designed to test the rhythmic performance of the player, encompassed in a gaming challenge. It was also designed to test different sound design principles. The first feature is a simple interactive method for the player to trigger sound while simultaneously shooting a laser beam. The second feature is an in-game challenge to complete. The first feature is necessary as without the ability to trigger sound, we would not be able to answer our question. The second feature is necessary to provide context for the first feature.

There were two different versions of the game used in the study. In each version, successfully shooting a block will result in a different sonic outcome. In version A, the sound is non-diegetic, and the player's interaction is performing part of the soundtrack; whereas in version B, the sound generated by the player's interaction is diegetic and is supplementary to the soundtrack. The impact of these two different outcomes will help inform future game design. If player experience is positively enhanced by using player interaction for non-diegetic sound design, further research should be done with an in-game challenge not explicitly tied to synchronised audio-visual outcome to isolate the strength of this effect. If player experience is negatively impacted, this may form strong enough evidence to sufficiently state that synchronised sound design should not interfere with non-diegetic sound design. The extent to which the playing experience correlates with the player's sense of immersion could also yield intrigue for further research. If players report a greater sense of immersion in one version over the other, we could undertake further research to ascertain why this occurs. All participants played both versions in a random order. The game moves multiple target blocks towards the player sequentially, in time with the music. In the 'A' version of the game, when players shot and destroyed a block, it would trigger a corresponding note (pitch and duration) for the bass synthesiser to play. Shooting the blocks in sequential order over the nearest playhead rectangle results in the note playing coordinated with the rest of the music. Without shooting the blocks, there would be no bassline in the music, only drums and a pad. In the 'B' version, the game music is the same and contains the same bassline, with the blocks in corresponding note order. However, the blocks and bassline notes are now synchronised from the furthermost playhead from the player, the back wall where blocks first enter the player's game environment. When the player shoots a block in this version, a sound effect is triggered. If the players fail to shoot the target blocks before they reach the player, they then take damage. This triggers a sound effect, along with a visual pulsing of red within the player's field of vision. This gets more intense as the player takes more damage. If the players take too much damage, the game stops, and the player has lost.

The primary aim of the game is to shoot all the blocks and make it to the end of the level. The secondary aim is shooting all the blocks in time with the music so that each of the target blocks plays the correct note at the right time. The difference between the two versions means there is significantly more demand for the player to shoot the target blocks accurately in Game A, as the player's interaction triggers a major part of the non-diegetic soundtrack. In the 'B' version of the game, the timing of the target block being destroyed has no impact on the non-diegetic soundtrack. This tests the impact of synchronicity in audio-visual game challenges and the impact of different sound-design choices. The game was developed using JavaScript with the MaxiInstruments library for audio events and the A-frame library for 3D graphics. This method was used after the successful completion of the previous study on The Impact Player Agency Over Rhythmic Sequence and Platform Placement on The Playing Experience (found on page 98), as the approach proved successful at yielding meaningful results. Additionally, this was chosen over the previous use of Unity from the Preliminary Study as a response to the COVID-19 outbreak. Developing specifically for a low overhead web platform on a commercially available headset (Oculus Quest) allowed for easier delivery and surveying of participants, as opposed to a fixed location with an Oculus Rift Dev kit and PC.

Game Design and Dynamics

The game introduces the concept of rhythm and musical performance as a typical music game challenge. The challenge of the game is to destroy all the blocks moving towards the player without the blocks hitting them. To destroy the blocks, the players can shoot lasers out of their respective left and right Oculus touch controllers with the press of a trigger on the controller. There are two different colours of blocks; each can only be destroyed by a laser of the corresponding colour. The left controller shoots a pink laser, and the right controller shoots a yellow laser. The player is free to move their head around to orientate the camera. The players are instructed to destroy all blocks and make it to the end of the level. If the blocks collide with the player, the player will take damage. After sustaining multiple instances of damage, the game will be rendered incomplete, and play will stop. The blocks are generated based on the midi arrangement of the bass synthesiser from the backing music, which was arranged by this researcher. The sound design for the two different versions of the game is different. In the ‘A’ version of the game, the backing music consists of a drumbeat that develops over time and a choral pad with the bass synthesiser missing. When the player correctly shoots and destroys a block, it triggers the corresponding note (pitch and duration) from the arrangement that the block represents at that exact moment in time. In the ‘B’ version, the backing music consists of the same drumbeat, choral pad, and bass synthesiser. When the player destroys the block, a sound effect is played. There is also another difference between the two versions regarding block spatialisation. In the ‘A’ version of the game, the spatial reference point of any note being triggered (in order to be synchronised with the original predetermined arrangement) is marked by a grey/white line that increases in alpha as the block approaches. This can be seen in Figure 36 below. There are four of these lines (they can be considered playheads) representing a beat in a bar of the metric arrangement. This can be clearly explained visually, in Figure 36 we can see the line is grey as the pink block is approaching the playhead but is not directly over it. In Figure 37 the yellow block is nearly fully over it so the playhead light is a lighter grey. In Figure 38 the lower yellow block is directly over the line so the playhead line is now white. The closest line to the player is where maximum synchronisation occurs. In the ‘B’ version, the spatial reference point of any note being triggered is the furthest playhead away from the player, which is displayed as a white semi-transparent wall opposite the player.

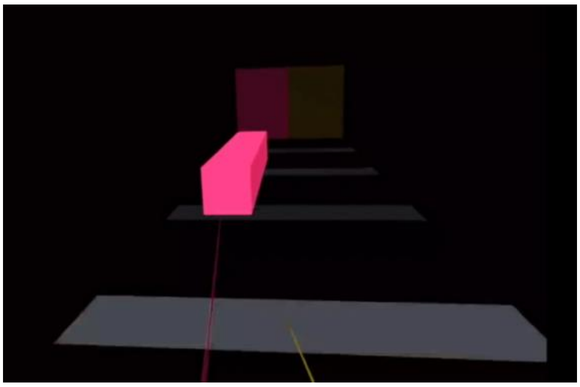


Figure 36: Grail: A pink target block approach the nearest playhead to the player

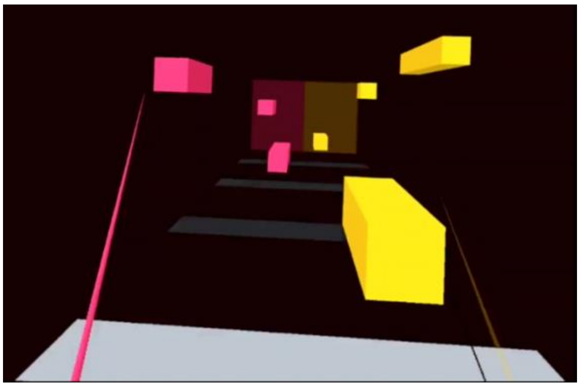


Figure 37: Grail: Multiple targets now visible ready for the player to shoot

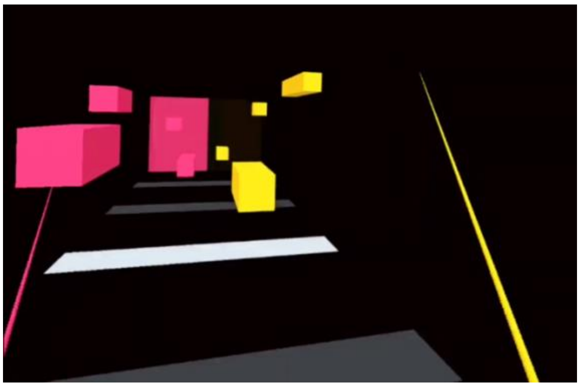


Figure 38: Grail: A pink target has spawned a new block. One target is a beat away from the optimal sync point

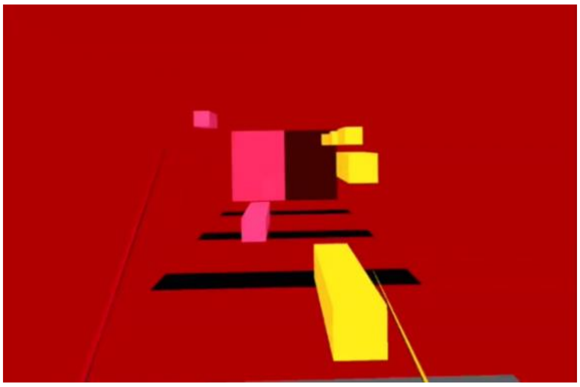


Figure 39: Grail: The player has taken lots of damage!

As mentioned above, when the blocks hit the player, the player will take damage. This causes the player's vision to be obscured by flashes of red that gradually become more opaque upon taking increasing amounts of damage, as seen in Figure 39. These flashes pulsate in time with the meter of the music. In both versions of the game, when the player takes damage, a sound effect is played. The game finishes if the player makes it to the end of the music or if the player takes too much damage. The latter stops the music and leaves the player with the vision of the scene fully obscured by red. In either of the versions, the player is not explicitly told to synchronise their input with the music. They are not punished for triggering the sound too early or too late in reference to the backing music and the original arrangement of the bassline. They are only punished for missing a target block. This was a deliberate design decision to test the playing experience regarding intuition, synchronised input, and immersion. As the number of blocks is fixed to the given level arrangement, there is a fixed maximum amount of playtime. Players were encouraged to play as many times as they wanted to, grasping a firm understanding of how the game played.

Once the game was developed, it was provisionally tested for bugs and performance-related issues, receiving some optimisation for performance for the Oculus browser. A splash webpage was created detailing the required game information for the study, with a randomised link to the first version of the game and the necessary GDPR details for online data collection (Goldsmiths, University of London, 2023). The link to the version of the game was randomised to avoid any potential bias arising out of the playing order of the two versions of the games - there is potential for players to find the game easier or more intuitive once they have gained playing experience. As the aim was to recruit participants online, it was accepted they would likely be using their own headset and subsequently would be familiar with the basic functionality of the device. Consequently, there was no need to introduce them to the device. Participants who were recruited locally were asked if they had any prior experience with the headset and given some time to gain familiarity with the controls before starting. As the camera position within the game was fixed, there was no need for the players to move around. As a result, there was no need to introduce the players to the 'guardian player protection system' designed by Oculus for player safety.

Participants gave informed consent by clicking the link and submitting the survey. They were informed that they were free to stop playing and leave the study at any point, should they wish and were given details on data collection and processing. Once the link was clicked, they were taken to the next webpage. This contained the A-frame canvas with the game, a disclaimer that there was flashing imagery and it was not recommended to play if they were epileptic or sensitive to light, a description detailing controls for the game, instructions for how the game is played, including the in-game challenge given, a link to the corresponding survey to be completed after playing, and lastly the link to the second version they would be due to play. The players were instructed on how to play and how to start the game (enter full VR mode). They were free to play for as long as they wanted, using as many attempts as they wished (by refreshing the page), but encouraged to play for around five minutes. Each version of the game required the player to fill out a linked survey hosted on Google Forms, with the same survey questions asked for each version of the game. The website for the study was hosted on my GitHub pages account.

In order to recruit participants with the most relevant background, the study was promoted in specific targeted locations, such as the 'OculusQuest' and 'Virtual-Reality' subreddits, the Facebook Quest gaming forum, The Creative Computing and Games Programme mailing lists, and direct messaging former music alumni and colleagues known to have an interest in gaming, and on the researcher's personal social media accounts on Twitter and Instagram. Participants were offered a £5 Amazon voucher for proof of survey completion (the name given for both survey submissions photographed with their headset).

The first section of the survey was similar to both the preliminary pre-survey and the first section of the survey used in Study Two. It contained questions relating to demographic characteristics, experience with music performance and electronic music software, and experience playing video games. This can be found in the appendix on page 255. The first section of the survey contained nine questions written by this research. It contained specific questions about the game the participants had just played, such as 'To what extent were you focused on syncing your input to the backing music?' answering on a five-point Likert scale. The second section of the survey contained the same established (and validated) immersion questionnaire (IEQ) (Jennett al., 2008). As discussed in the previous two study chapters on pages 76 and 106, the IEQ was used as it contained the most relevant questions for measuring player immersion in a gaming experience. As the 31 questions were used in both previous studies, this also allows a comparison between all three different game prototypes across each differing version.

Each survey group was exported from Google Forms in CSV format. They were combined in a single Excel sheet with an additional label variable added to identify which version of the game/group it was (with A or B). The questions relating to demographic characteristics and prior musical and gaming experience were separated from the survey data in Likert format. Additionally, the multiple-choice questions from the survey were also filtered out; 'Which sensory feedback was the most useful in guiding your input?'. The data from this question and the questions relating to demographic characteristics will be addressed outside of the

Paired-Sample testing separately. Lastly, the 10-point Likert scale question used as a baseline of the IEQ questionnaire was also removed from our analysis. This will be used separately when analysing the IEQ results. The remaining data is all on a five-point Likert scale. In a similar fashion to the previous Preliminary Study (page 76) and to the follow-up study (page 106), these remaining questions were split into three sections for analysis. One section comprised only the questions written by the researcher, one comprised of the 31 IEQ questions, and the third with both sets of questions combined. It is also worth noting for questions 6, 8, 9, 10, 18, and 20 of the IEQ, a negative scale was used, and the results for these were reversed before being exported (so a '1' = '5' and vice versa). After the data was cleaned, it was imported into SPSS.

For each section (and both sections combined) Cronbach's alpha was calculated to measure internal reliability. The section comprised of the eight questions by this research had a Cronbach's alpha of 0.626. Although this could be considered reasonably reliable (Hinton et al., 2004) as there are only a small number of questions, the internal consistency can be improved by comparing the Cronbach's alpha value against each individual question's value. It was found that the question 'How Challenging was the game?' did not correlate strongly with the scale overall; by removing this question, we could improve the internal consistency to 0.670. A potential reason why this question was found to be inconsistent could be because regardless of how intuitive, responsive, or synchronised the player's interaction felt, the player still found the games equally challenging (M = 3.83, SD = 0.879). The second section, comprised of the 31 IEQ questions, had a Cronbach's alpha of 0.838, showing great consistency. By comparing the Cronbach's alpha value against each individual question's value, the following seven questions could be removed as they were found to be redundant. Subsequently, this improves the questionnaire's reliability (as described on page 57) (Koçak et al., 2014) and the alpha value increases (across each measure taken):

Q15: To what extent did you feel consciously aware being in the real world whilst playing?
Q17: To what extent were you aware of yourself in your surroundings?
Q18: To what extent did you notice events taking place around you?
Q19: Did you feel the urge at any point to stop playing and see what was happening around you?

Q26: To what extent did you find the game challenging?

Q29: To what extent did you find the game easy?

Q31: How well do you think you performed in the game?

After removing these questions, Cronbach's alpha for the remaining 24 questions was 0.895, which shows excellent internal consistency. As previously discussed in the Preliminary Study (page 77) and the follow-up study (page 107) Cronbach's alpha could be high as a result of redundant questions (Tavakol & Dennick, 2011). However, as argued in the previous study on page 99, given the participants are wearing a headset and playing a VR game, the first three questions are the most redundant questions for this survey. While the question 'Did you feel the urge at any point to stop playing and see what was happening around you?' could infer player engagement, I would consider it redundant due to the novelty and short duration of expected playing time. Local participants played in a quiet room with no distractions, and one would anticipate most remote VR users would likely play in a similarly quiet space, as they were aware they were going to play a music game. The remaining three questions that were removed, all related to the game challenge and player performance, potentially highlight the difficulty of the challenge. Most participants had gaming or musical experience but had little experience performing in VR.

The Cronbach's alpha for the combined researcher's questions and IEQ is 0.854, showing great consistency. This is optimised to 0.901 when the inconsistent questions discussed above are removed. The results from these 31 questions were then totalled and divided by the number of questions (31). Totals for each separate section of the survey were also generated. Section one (consisting of my questions) is the first eight questions, section two being the remaining 23 questions from the IEQ. The results of the complete 31-question IEQ will also be discussed relative to the questionnaire's model; this will be examined separately.

Data Analysis Combined Data

To determine the most appropriate test to use for comparing means, the distribution of the data must be checked for normality (Löfgren n.d.) (Field et al., 2013). Initially, this is performed by checking the histogram and Q Plot for the two separate groups (A and B versions of the game) by eye. The two histograms and Q Plots can be seen in the figures below.

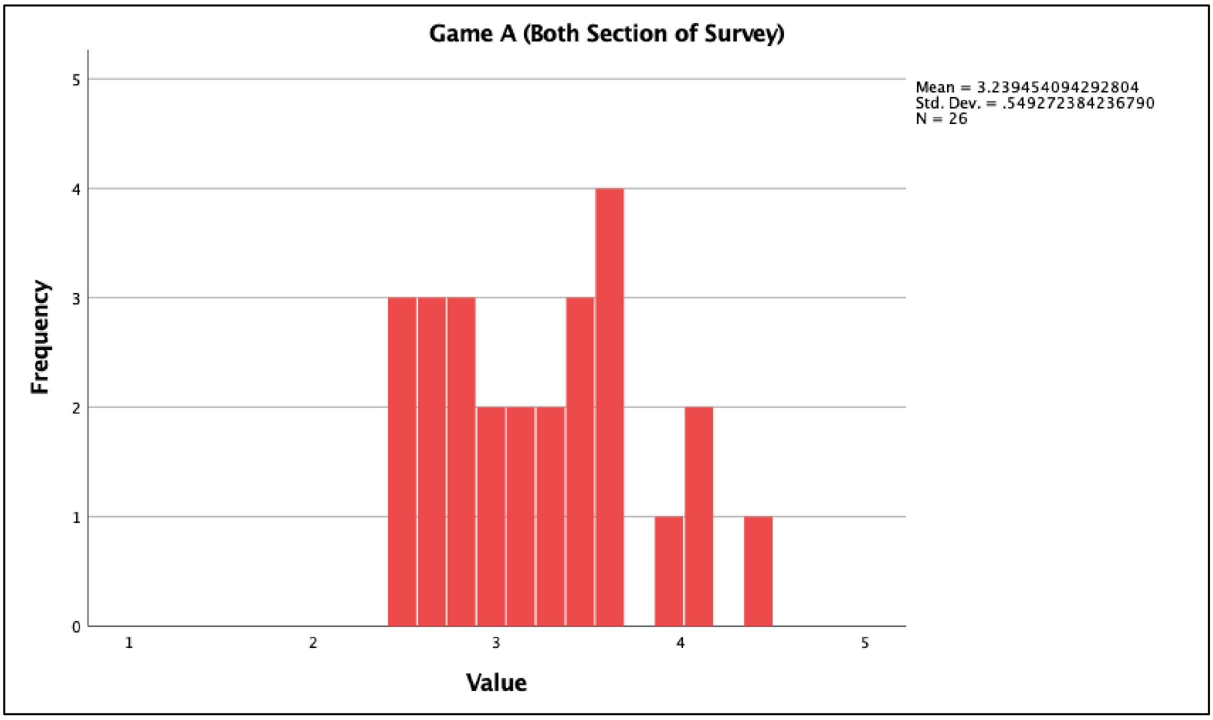


Figure 36: Study Three Game A Histogram Results for Complete Survey

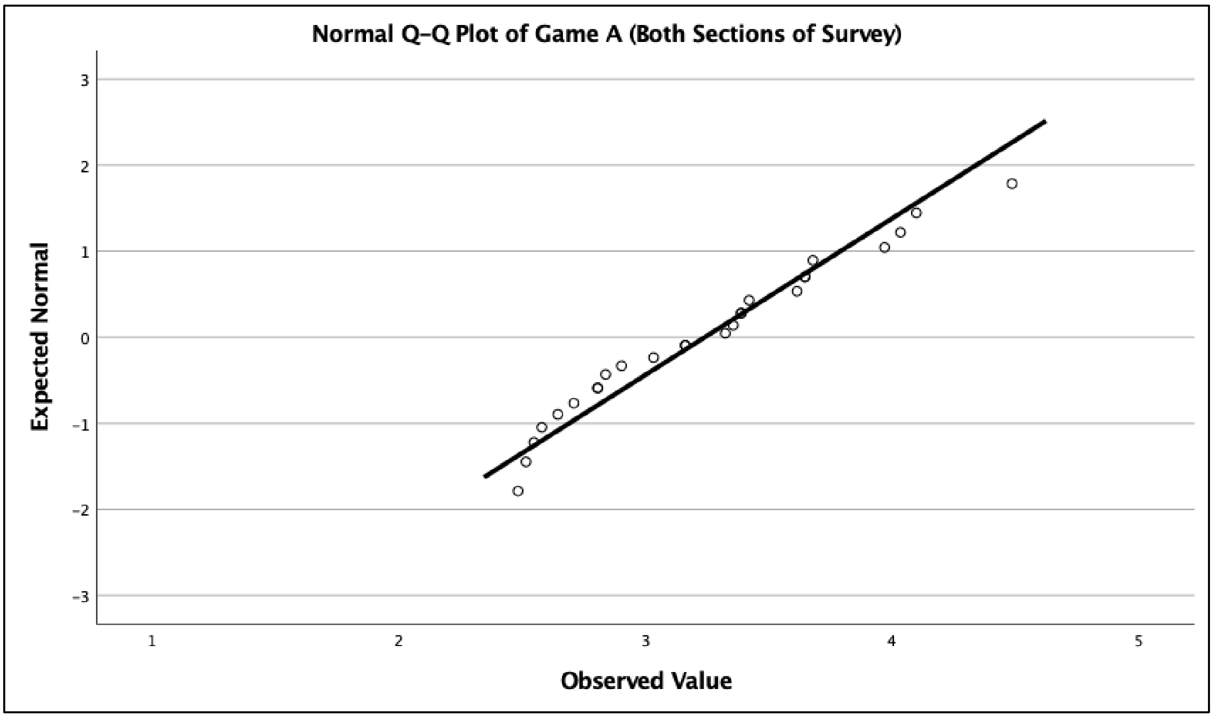


Figure 37: Study Three Game A Q-Plot for Complete Survey

In the distribution of the histogram for Game A (the version where players triggered a bass note in the arrangement of the music by shooting the target blocks), the distribution is skewed below a mean of 3.239, with a roughly normal distribution. By calculating the Z-values for Skewness (S = 0.436, SE = 0.456) and Kurtosis (S = -0.570, SE = 0.887), we find both to be within the acceptable range of -1.96 and 1.96 for normality (Löfgren n.d.). This is further supported by the group of values shown in the corresponding Q-plot above, as there is a visible trend in values. The distribution visible in the histogram for Game B (the version where players trigger a sound effect upon hitting a target block and where the bassline is present in the backing music) in the figure below has comparatively more visible normalcy by eye, around a mean of 3.242.

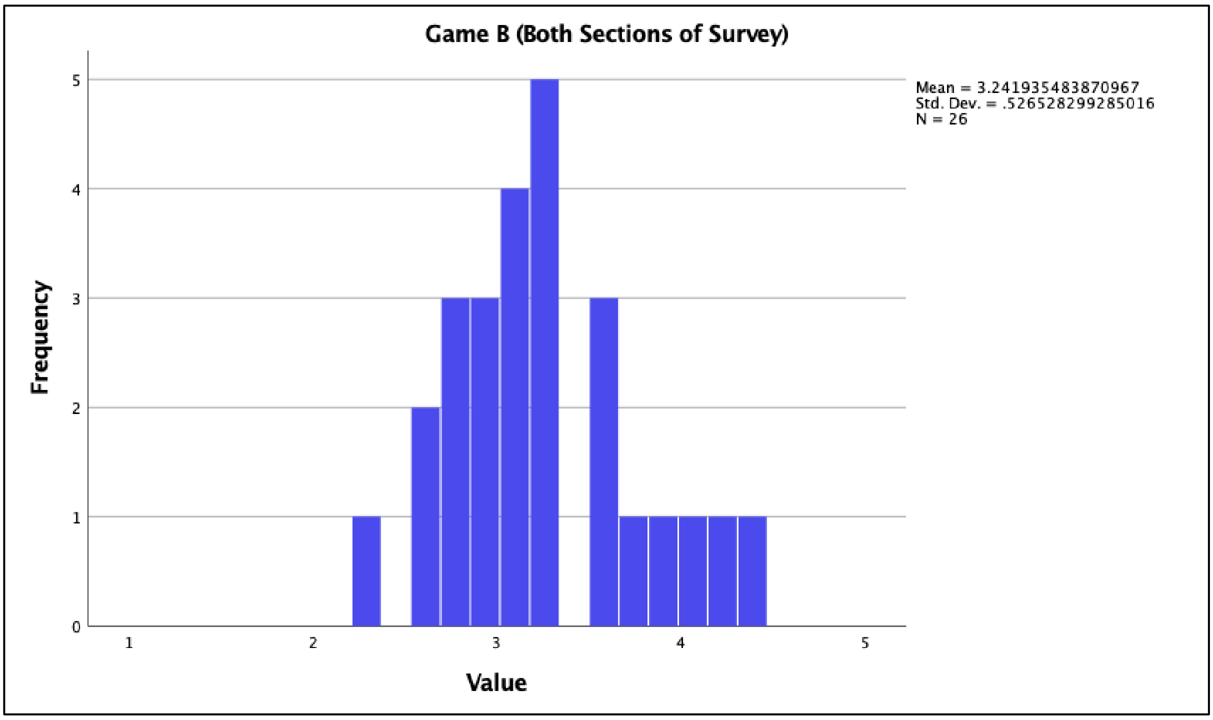


Figure 38: Study Three Game B Histogram Results for Complete Survey

The results were also checked by calculating Z-values for Skewness (S = 0.461, SE = 0.456) and Kurtosis (S = -0.219, SE = 0.887); both were found to be within the acceptable range (-1.96 to 1.96) for normality (Löfgren n.d.). This is further supported by the Q-Plot for Game B in the figure below; once again, the trend is visible to the eye.

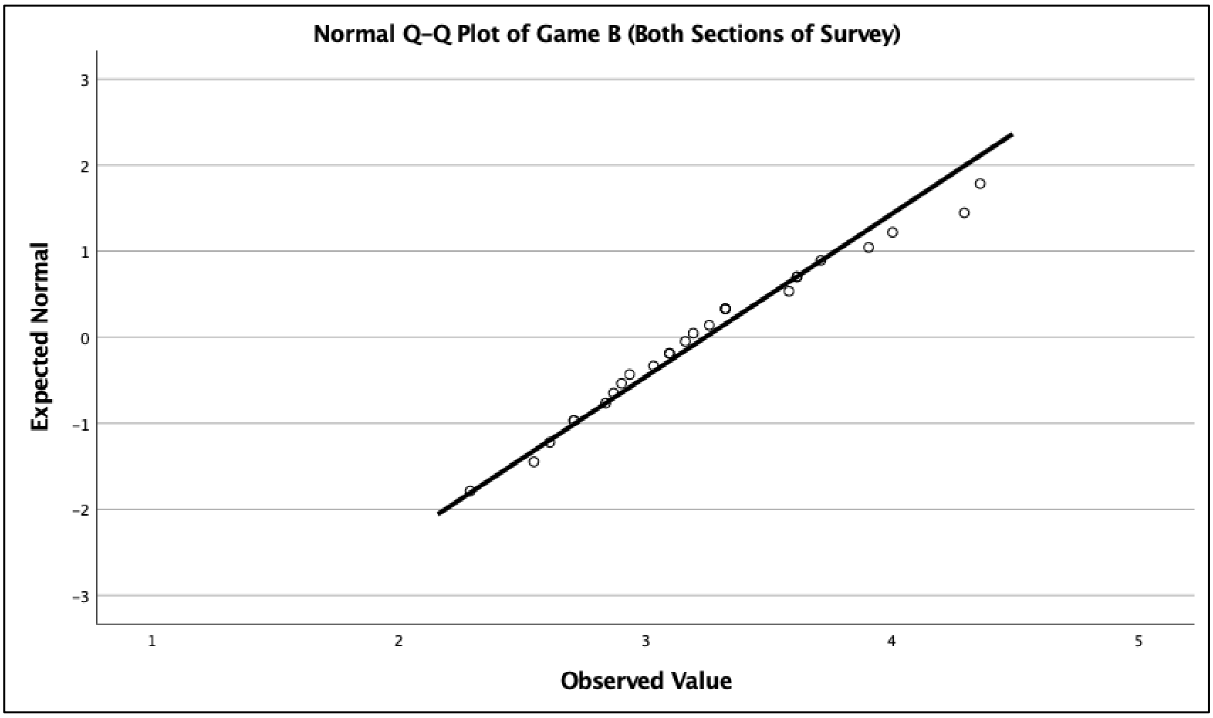


Figure 39: Study Three Game B Q-Plot for Complete Survey

As the sample size is small, a Shapiro-Wilk test ($p < 0.05$) (Shapiro & Wilk, 1965) of normality was performed, which also showed the null hypothesis to retain the 0.05 level of significance, therefore assuming normality for the dataset (Game A = 0.795, Game B = 0.328). As both groups contain the same participants, use the same surveys, and have normally distributed results, it has met the necessary assumptions (Field et al., 2013) for the Dependent-means/Paired-Sample T-test to be the most appropriate method of comparing means between the playing experience of both versions.

Table 24: Paired Sample Statistics for 3rd Study Survey (Combined)

Paired Sample Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair	GABoth	3.239454094292804	26	.549272384236790	.107721177134655
	GBBoth	3.241935483870969	26	.526528299285016	.103260695096658

The label 'GABoth' means the pair contains the results from both sections of the survey: Section One, written by this researcher, and Section Two contains the IEQ for the Game A version, whereas 'GABBoth' contains the results from the same corresponding sections for the Game B Version.

Table 25: Paired Samples Correlations for 3rd Study Survey (Combined)

Paired Samples Correlations				
		N	Correlation	Sig.
Pair	GABoth & GBBoth	26	.765	.000

The results show a statistically significant result of 0.000, meaning the difference is not due to chance.

Table 26: Paired Samples Test for 3rd Study Survey (Combined)

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair	GABoth–GBBoth	-.002481389578165	.369708727709088	.072505846805162	-.151809976373408	.146847197217079	-.034	25	.973

Cohen's d was used to calculate the effect size (Wiseheart n.d) (Cohen, 1988). The results show the players had a difference in playing experience and sense of immersion between the two different versions of the game. The results show players favoured the 'B' version of the game (M= -0.002, SE= 0.370, $t(25) = -0.034$, $p < 0.05$, with an effect size $d = 0.370$), although the reported difference between the two versions is small. To get a better understanding of where this difference occurs, we can compare this result to each individual of the section and how it relates to each group.

Data Analysis: Section One

The distribution for the seven questions written by this researcher can be found in the figures below.

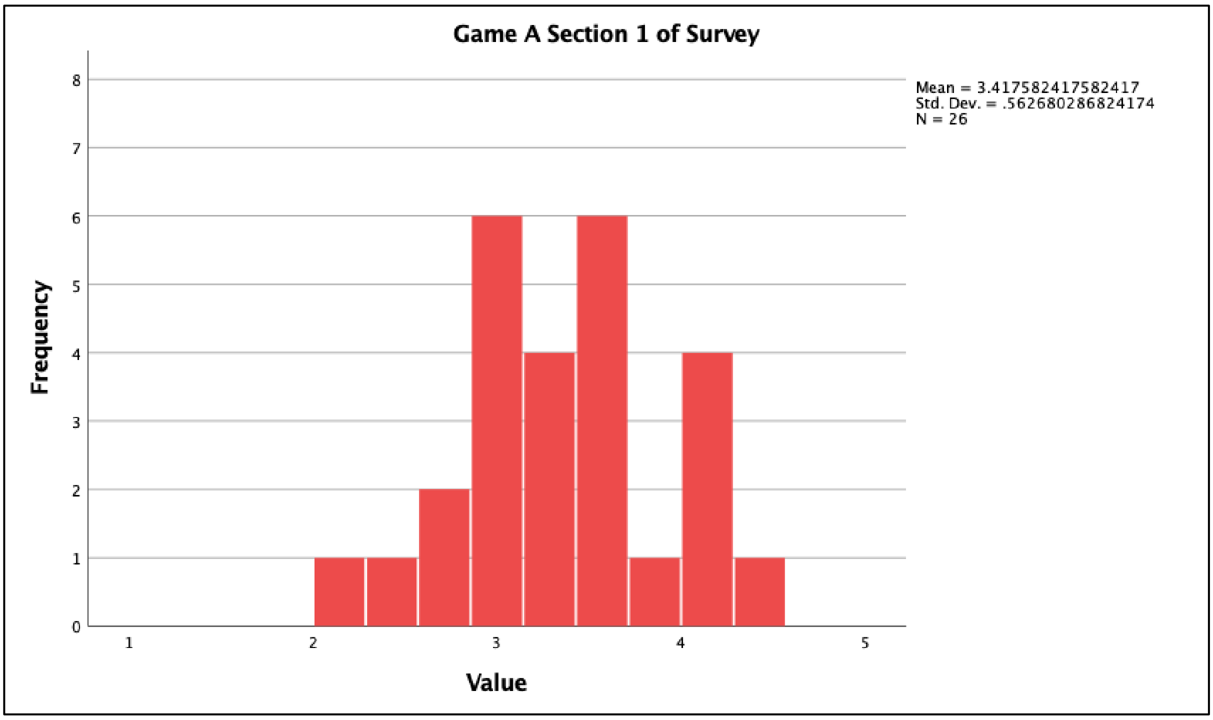


Figure 40: Study Three Game A Histogram Results for Section One of the Survey

In the figure above, we see the distribution for section one of Game A. As approximated by the eye, these results look closer to an evenly normal distribution compared to the overall results. They are also concentrated around a slightly higher mean of 3.418 compared to the combined Game A results (M = 3.239). This is supported by the Q-Plot in the figure below.

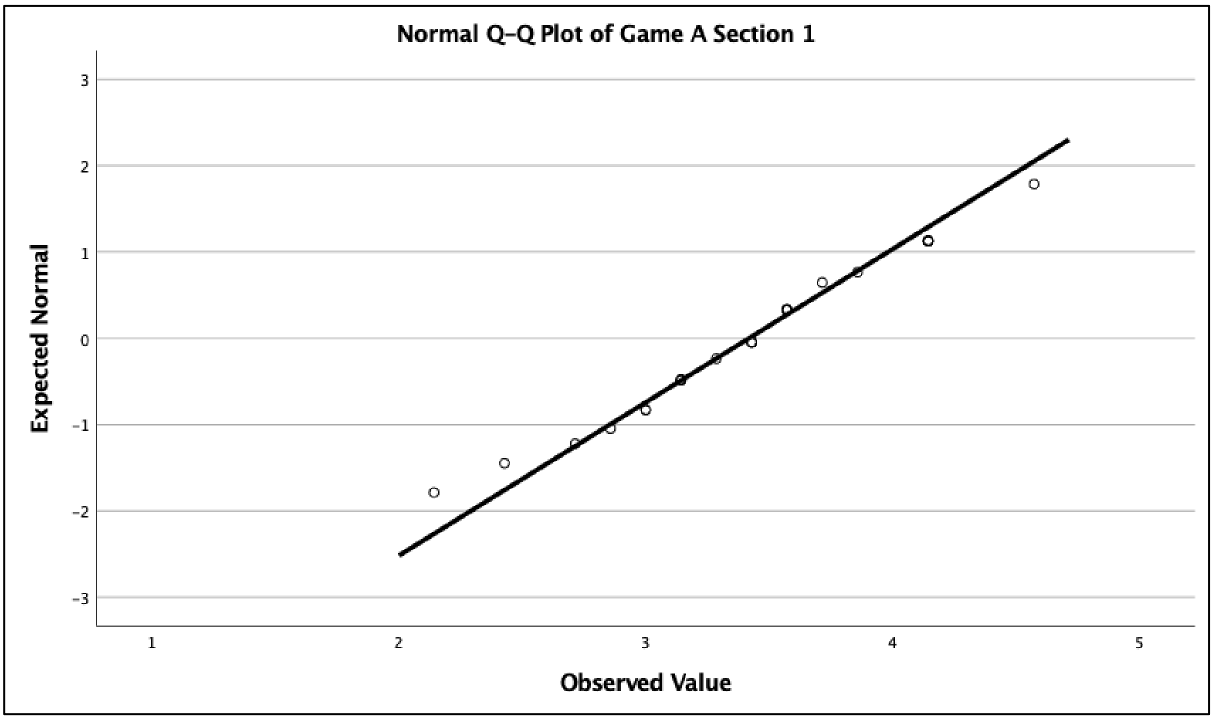


Figure 41 : Study Three Game A Q-Plot for Section One of the Survey

The distribution for the Game B results of section one of the survey can be seen below.

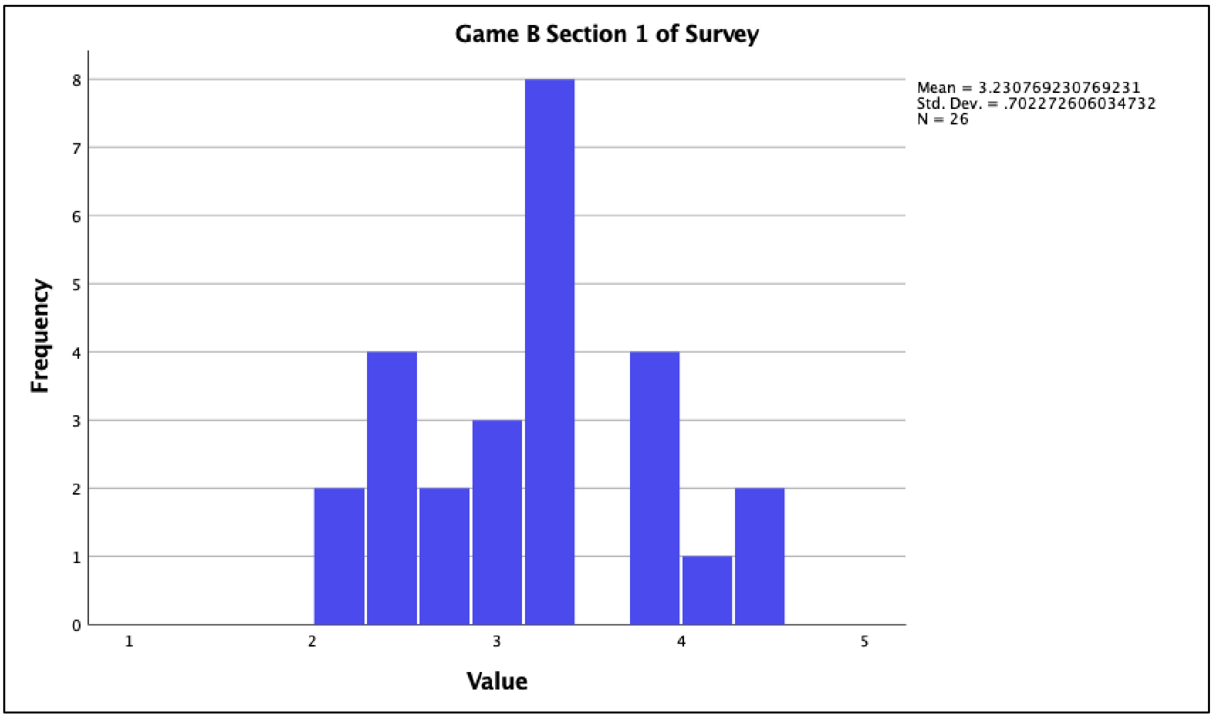


Figure 42: Study Three Game B Histogram Results for Section One of the Survey

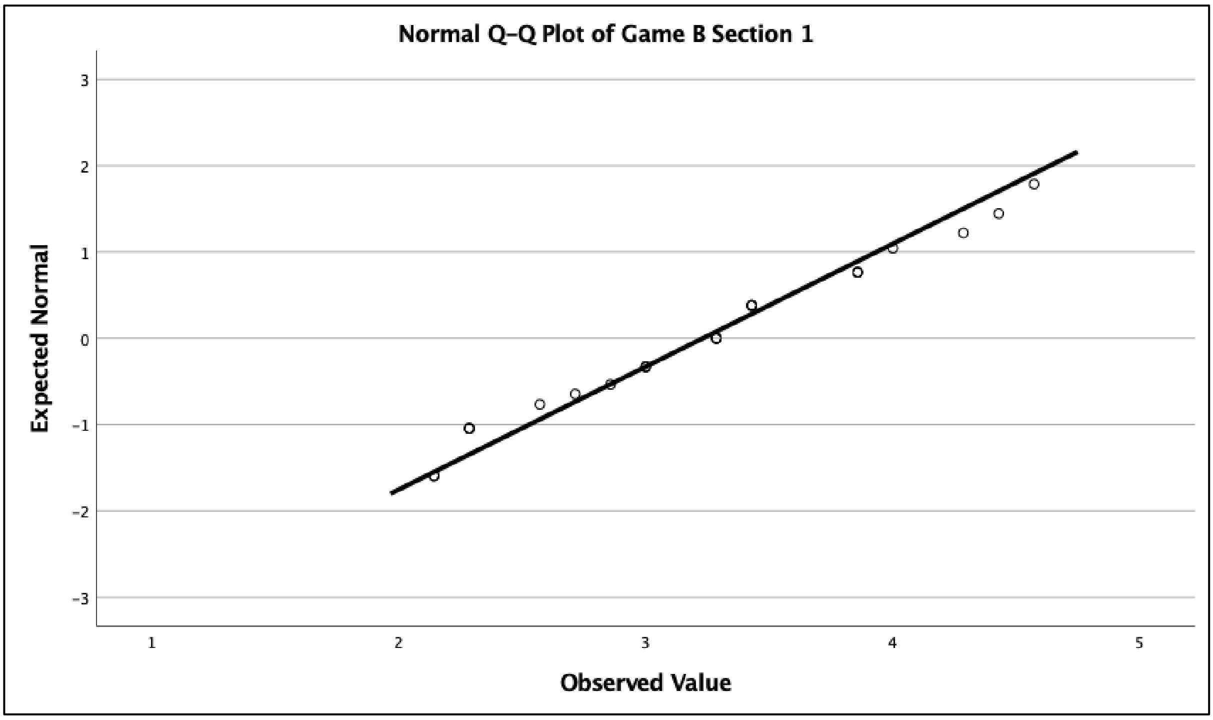


Figure 43: Study Three Game B Q-Plot for Section One of the Survey

The distribution, as approximated by the eye, also appears to be more visibly normally distributed when compared to the combined results. However, we see the values concentrated in a narrower band of values than the combined B results, with a slightly lower mean of 3.231 than that of the combined Game B Results (M = 3.242). To confirm normality the Z-values for Skewness for both Game A (S = -0.146, SE = 0.456) and Kurtosis (S = 0.131, SE = 0.887) and Game B Skewness (S = 0.113, SE = 0.456) and Kurtosis (S = -0.714, SE = 0.887) this was confirmed and found to be in an acceptable range of normality (Löfgren n.d.). Using the Shapiro-Wilk test, we have shown the null hypothesis retains the 0.05 level of significance, therefore assuming normality (Game A p= 0.795, Game B p= 0.328). Subsequently, we can again use a Paired-Sample T-test to compare means:

Table 27: Paired Sample Statistics for 3rd Study Survey (Section One)

Paired Sample Statistics					
		Mean	**N**	**Std. Deviation**	**Std. Error Mean**
Pair	GA1	3.417582417582418	26	.562680286824173	.110350683170402
	GB1	3.230769230769231	26	.702272606034731	.137726989308952

The label 'GA1' means the pair contains the results for only the first section of the survey for the version of the game that triggered the bass note in the music arrangement upon destruction of the target block. Whereas the label 'GB1' contains the first section of the survey for the version where the backing track contains the full bassline and the player triggers a diegetic sound effect.

Table 28: Paired Samples Correlations for 3rd Study Survey (Section One)

Paired Samples Correlations				
		N	**Correlation**	**Sig.**
Pair	GA1 & GB1	26	.445	.023

Table 29: Paired Samples Test for 3rd Study Survey (Section One)

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair	GA1 –GB1	.186813186813187	.677051507725247	.132780724984556	-.086653835355018	.460280208981392	1.407	25	.172

The results show a statistically significant result of 0.023, meaning the difference is not due to chance. The effect size was calculated using Cohen's d (Wiseheart n.d) (Cohen, 1988). Consequently, on average, participants players reported that in version A, the audio-visual experience felt more synchronised, that it felt more like a musical experience, was intuitive, was responsive, that they focused on syncing their input to the backing music, and that by making mistakes it broke their focus more ($M = 0.187$, $SE = 0.677$, $t(25) = 1.407$, $p < 0.05$, with an effect size $d = 0.677$). This is contrary to the analysis of the combined sections of the survey. It is not surprising that players responded that version A was more like a musical experience, given they are triggering sonic events with pitch and duration as opposed to the sound effect in version B. Interestingly, this could potentially impact the perceived intuitiveness or responsiveness of the interaction, given that the method of interaction and the in-game challenge are identical between the two versions. Players also reported, that when their actions generated non-diegetic sound to form part of the backing music, they focused more on synchronising their playing inputs. This is surprising given they have less 'complete' backing music to synchronise their actions to, especially as players are only informed to destroy blocks and not explicitly told to synchronise their input.

The result of the multiple-choice question from the first section, 'What sensory feedback was the most useful in guiding your input?' is also useful to inform our understanding of the player's experience. In Game A, there was a 50% split of players (13 out of 26) reporting they used the audio cues / their ears to guide their input, as opposed to using the visual cues of the flashing playhead lines and position of the target blocks. In Game B, only 23.08% (6 out of 26) of players reported using the sound or their ears to guide their input. This supports the findings from the Paired-Samples test, although it is surprising. As in Game B, there is an existing one-to-one audio-visual relationship between the 3D graphics that exist in the game environment and what the player is hearing at any one point. One (wrongly) could expect that this would serve as a more useful guide for the player's cognition. The evidence from these findings suggests that giving players more responsibility for their sonic experience in-game by impacting the non-diegetic sound in the game leads to a more engaging experience. This is especially important if this is coupled with a higher consequence on their focus if they make a mistake; this will be discussed further in the survey evaluation after the remaining analysis. We should also consider this in relation to our demographic, which will be analysed after section two of the survey.

Data Analysis: Section Two

The distribution for the remaining 24 questions related to player immersion was checked for normality as previously described. The analysis is below:

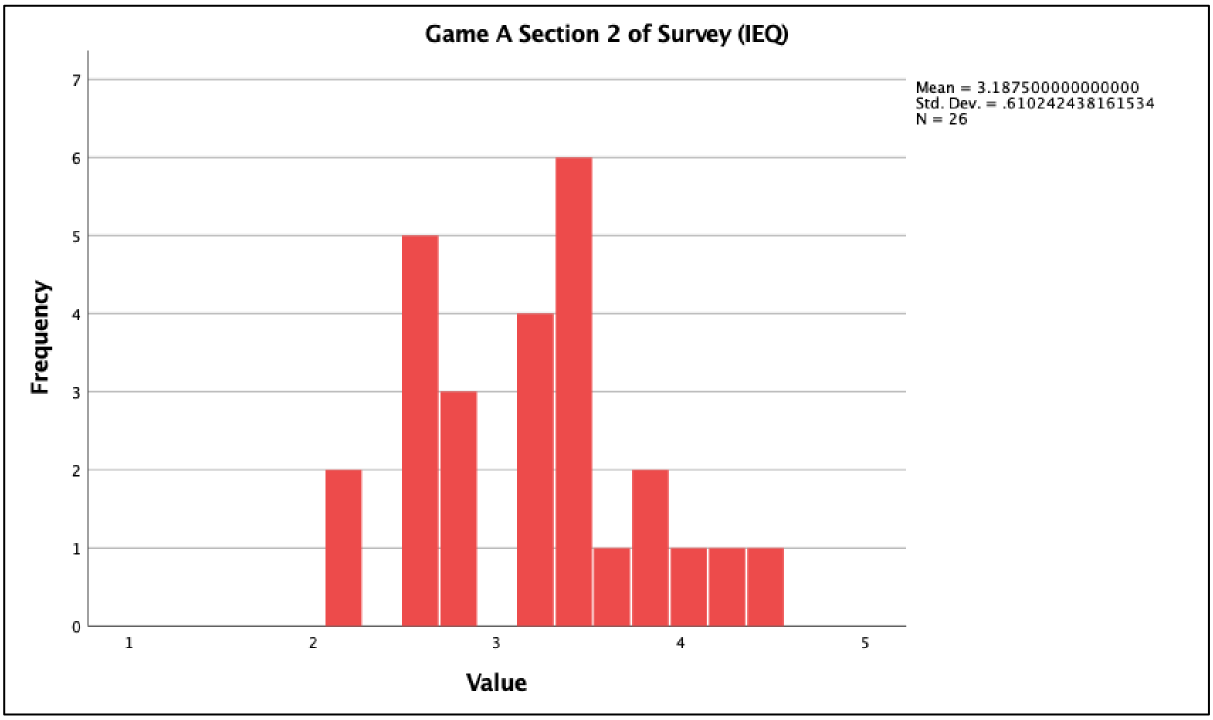


Figure 44: Study Three Game A Histogram Results for Section Two of the Survey

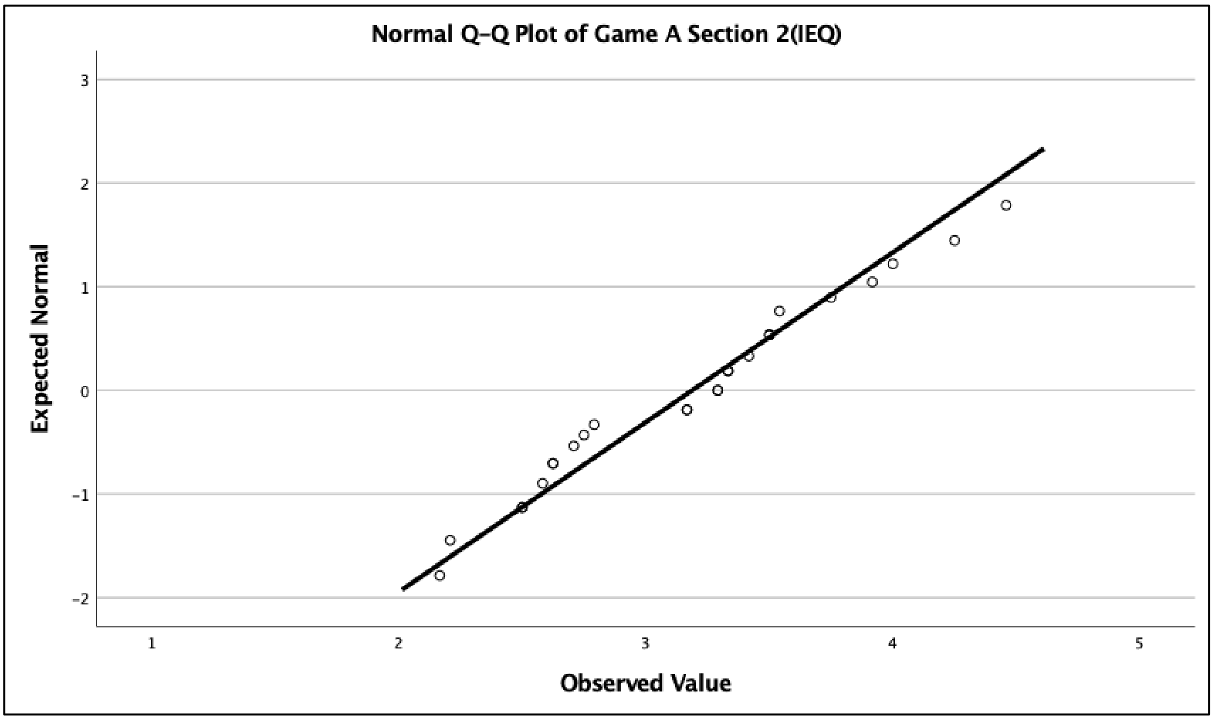


Figure 45: Study Three Game A Q-Plot for Section Two of the Survey

The distribution of the Game A section two results, as approximated by the eye, looks roughly normal and similar to the distribution of the combined results, although more skewed compared to the equivalent section one. After checking the Z-values for Skewness (S = 0.196, SE = 0.456) and Kurtosis (S = -0.592, SE = 0.887), this was confirmed and found to be within an acceptable range for normality (Löfgren n.d.). The distribution of Game B is below.

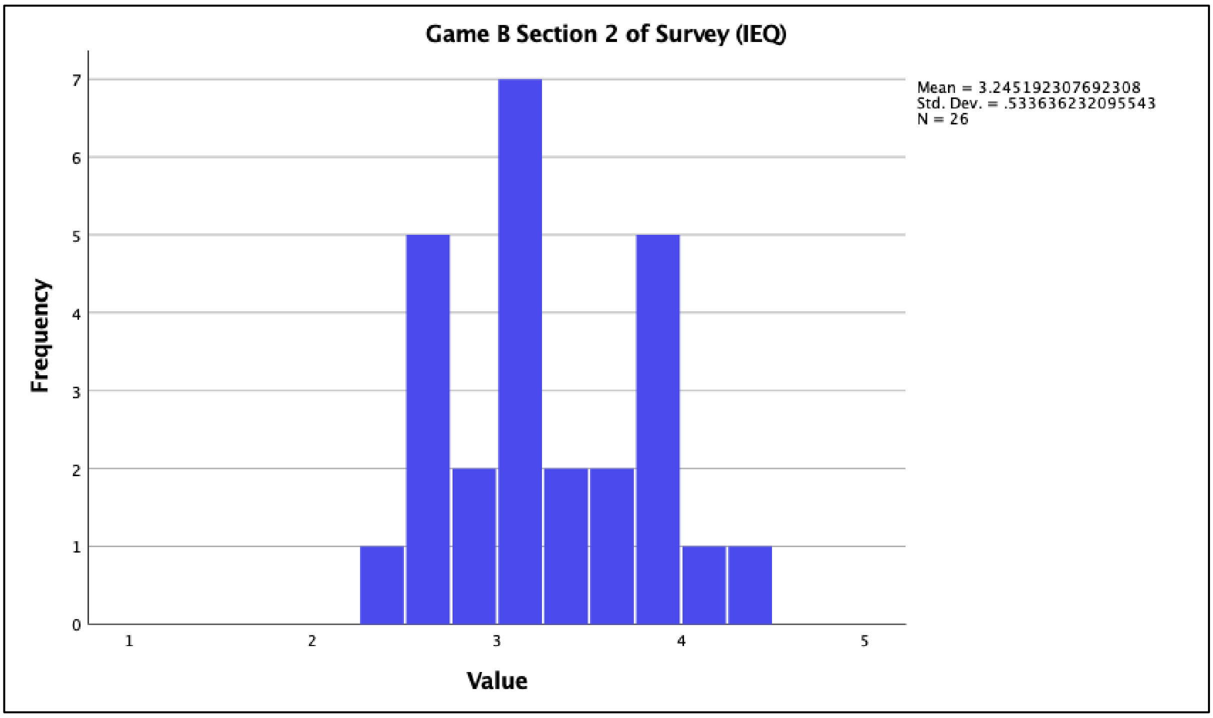


Figure 46: Study Three Game B Histogram Results for Section Two of the Survey

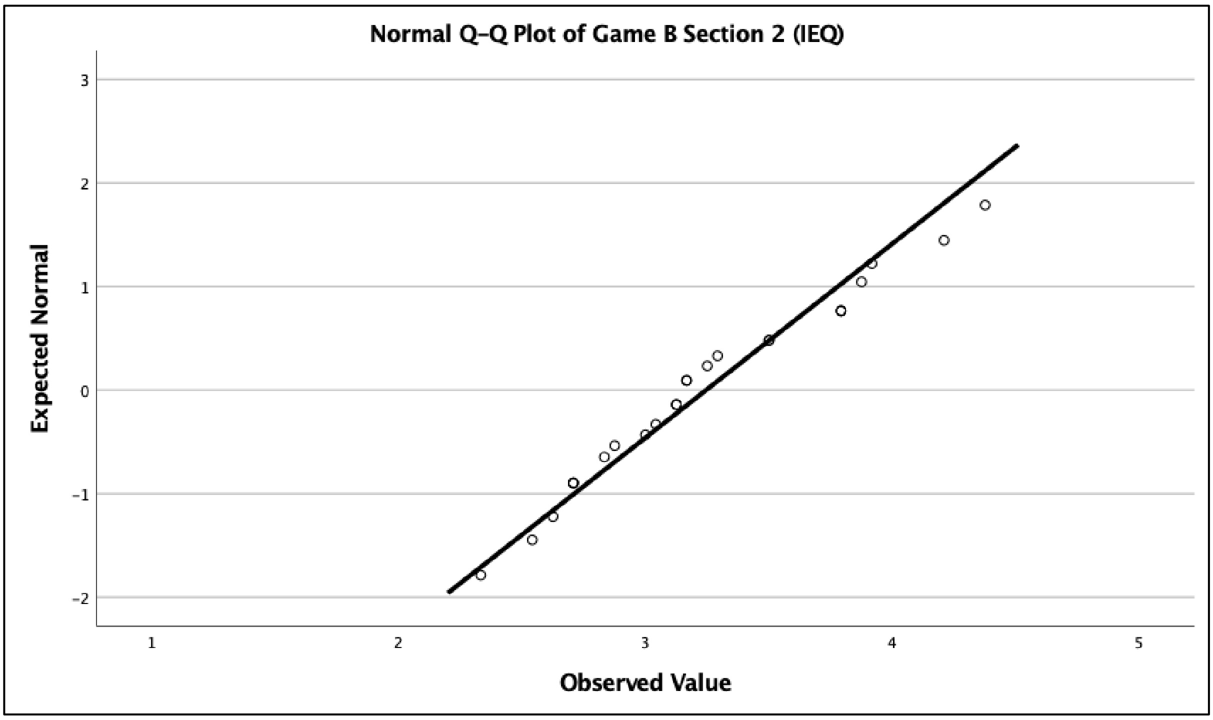


Figure 47: Study Three Game B Q-Plot Results for Section Two of the Survey

The distribution of section two for Game B was approximated by the eye and found to be normal, albeit slightly skewed. The distribution was compared to both the combined results and the results from section one. This was confirmed after checking the Z-values for Skewness (S = 0.414, SE = 0.456) and Kurtosis (S = -0.565, SE= 0.887) and found to be within the acceptable range for normality (Löfgren n.d.). The Shapiro-Wilk test also showed the null hypothesis to retain the 0.05 level of significance, therefore assuming normality (Game A p = 0.510, Game B p = 0.441). Subsequently, a Paired-Sampled t test was once again used to compare means.

Table 30: Paired Sample Statistics for 3rd Study Survey (Section Two)

Paired Sample Statistics					
		Mean	**N**	**Std. Deviation**	**Std. Error Mean**
Pair	GA2	3.187500000000000	26	.610242438161534	.119678388469542
	GB2	3.245192307692308	26	.533636232095543	.104654675408281

The label 'GA2' means the pair contains the results for only section two of the survey for Game A, the version of the game where players trigger the bass note in the music arrangement by shooting the target blocks. Whereas the label 'GB2' contains the results only for section two of the survey for Game B, the version where the backing music already contained the bassline, and players triggered a sound effect by destroying the target blocks.

Table 31: Paired Samples Correlations for 3rd Study Survey (Section Two)

Paired Samples Correlations				
		N	**Correlation**	**Sig.**
Pair	GA2 & GB2	26	.782	.000

The results show a statistically significant result of 0.000, meaning the difference is not due to chance. The effect size was calculated using Cohen's d (Wiseheart n.d) (Cohen, 1988).

Table 32: Paired Samples Test for 3rd Study Survey (Section Two)

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	**Std. Deviation**	**Std. Error Mean**	**Lower**	**Upper**	**t**	**df**	**Sig. (2-tailed)**
Pair	GA2 -GB2	-.057692307692307	.384251965056228	.075358010306080	-.212895035176448	.097510419791833	-.766	25	.451

On average, participants had a greater sense of immersion when playing Game B, the version of the game where the backing music contains the bassline that corresponds to the target blocks and where the player's target shooting is sonified with a sound effect. Compared to playing Game B, where the player triggers the bass notes in the arrangement ($M = -0.058$, $SE = 0.384$, $t(25) = -0.766$, $p < 0.05$, with an effect size $d = 0.384$). This difference in the sense of immersion is quite small, and although this section of the survey is best understood as a whole, we can study some of the results that have a greater deviation from the mean individually to gain further insight. In question 3: 'How much effort did you put into playing the game?' players reported putting noticeably more effort into playing Game A ($M = 4.269$, $SD = 0.667$) compared to that of Game B ($M = 3.808$, $SD = 0.749$). Question 9: 'To what extent did you feel that the game was something you were experiencing, rather than something you were just doing?' players were more likely to report it as an experience than an action (something they were doing) in Game B ($M = 3.548$, $SD = 1.029$) compared to Game A ($M = 2.962$, $SD = 1.076$). Question 13: 'Were there any times during the game in which you wanted to give up?' players reported they were less likely to give up playing Game A ($M = 3$, $SD = 1.2$) compared to Game B ($M = 3.577$, $SD = 1.331$). Question 15: 'To what extent did you feel like you were making progress towards the end of the game?' players reported they felt they were more likely making progress in Game B ($M = 3.346$, $SD = 0.977$) as opposed to Game A ($M = 2.654$, $SD = 1.093$). These specific results will be addressed further in the following evaluation.

Data Analysis: IEQ Scoring

The original 32 IEQ questions (31 five-point Likert and one 10-point baseline Likert) provide scores for five factors of immersion: 'Cognitive Involvement', 'Emotional Involvement', 'Real-World Dissociation', 'Control', and 'Challenge' (Jennett et al., 2008). As previously described in the Preliminary Study Chapter on page 90 and in the Study Two Chapter on page 121, the totals for each factor have been normalised to the one-to-five scale of the Likert scale. This allows for direct comparison between factors as they are appropriately scaled. The individual participant IEQ scores from the first 31 questions for both versions of the game surveys were checked against their baseline, with most found highly comparable. It is not unexpected that some reported baseline values would be over or under the respective IEQ score, given it is a single 10-point Likert value that is normalised (divided by two) for the comparison. If participants incorrectly interpret the term 'immersed' to not factor in their real-world environment, as opposed to only their sense within the game environment, a single reported value may be incomplete. We can communicate this further by checking the total IEQ mean for Game A (M = 3.315, SD = 0.455) against the normalised baseline mean of Game A (M = 2.942, SD = 1.00), along with the IEQ score for Game B (M = 3.342, SD = 0.405) against the corresponding baseline of Game B (M = 3.404, SD = 0.837).

The following table contains the breakdown per immersion factor for Game A, with the baseline mean normalised to the five-point Likert scale.

Table 33: IEQ Scoring Immersion Factors for Game A (Study Three)

IEQ Scoring Immersion Factors							
	Cognitive involvement	Emotional Involvement	Real World Dissociation	Control	Challenge	Overall	Baseline
Mean	3.449	2.712	3.748	3.077	3.462	3.315	2.942
Std. Deviation	0.601	0.938	0.712	0.506	0.416	0.455	1.00

The table below contains the breakdown per immersion factor Game B, with the bassline normalised as above.

Table 34: IEQ Scoring Immersion Factors for Game B (Study Three)

IEQ Scoring Immersion Factors							
	Cognitive involvement	Emotional Involvement	Real World Dissociation	Control	Challenge	Overall	Baseline
Mean	3.568	2.795	3.736	3.131	3.231	3.343	3.404
Std. Deviation	0.472	0.876	0.624	0.527	0.412	0.405	0.837

We can see in the tables above that while participant baselines were under-reported compared to the IEQ score in the results for Game A and over-reported in comparison to Game B, their scores are still sufficiently close. As discussed above, this is not surprising as we can expect the outcome of 31 answers to be more detailed than a single baseline question. The three factors of Cognitive Involvement, Emotional Involvement, and Control reported an increase in immersion for Game B, whereas Real World Dissociation and Challenge reported an increase in Game A. However, the difference between each factor is comparatively small, with the largest reported difference in results being for Challenge (Game A M = 3.462, Game B M = 3.231). Interestingly, this difference is larger than the reported overall difference on player immersion when comparing a high-effort challenge (M = 3.66) against a low-effort challenge (M = 3.619), as reported in Cox's study 'Not Doing but Thinking: The Role of Challenge in the Gaming Experience' (Cox et al., 2012). It is also larger than the difference in immersion when comparing the low-expertise players (M = 3.725) against high-expertise players (M = 3.707) in low-effort challenges in the same study. The difference in the Challenge factor between the two versions is less than the difference found between low-expertise players (M = 3.967) versus high-expertise players (M = 3.361) in high-effort challenges (Cox et al., 2012). This is particularly interesting when related to the player's previous experience, where prior experience was found to have made a difference in reported immersion as discussed within this Chapter under the Demographics subheading below. The smallest reported difference was Real World Dissociation, which was also the highest reported factor of immersion for both versions. This is not surprising given the isolated nature of the VR headset, which covers the sight of your surroundings (as discussed in previous study chapters on pages 92 and 108). Giving the player a more musical, non-diegetic interaction in Game A led to a more immersive challenge (M = 3.462) compared to the diegetic sound effect used in Game B (M = 3.231). However, this method of sound interaction came at the cost of Cognitive Involvement, reflected in the results for this factor of player immersion (Game A M = 3.449, Game B M = 3.568). This is also reflected in the factor of Emotional Involvement, where players reported a slightly higher sense of Emotional Involvement (Game A M = 2.712, Game B M = 2.795). One could argue this occurs due to music affection and the impact of the player incorrectly synchronising their input to the intended bassline in Game A, compared to Game B, where the bassline is always correctly in time. The potential relationship between the factors and the rest of the results will be discussed in further detail in the evaluation below.

Demographics

The most important demographic questions asked were about prior gaming experience and musical experience. Of the 26 participants, 17 (65.38%) had played games for more than five years, with six (23.08%) playing for over a year but under five years. Two participants had experience of over a week but less than a month (7.69%), with a single participant (3.85%) having less than a single week's experience of playing games. Only two participants (7.69%) usually played games on an Oculus headset. The most common device participants usually played on was a gaming console (53.85%); only five participants (19.23%) reported usually playing with motion controls, with 13 participants (50%) reporting having never used motion controls before. The 13 participants who had used motion controls before found the interaction more intuitive (M = 4.154, SD = 0.674) than the 13 participants who had not used the motion controls before (M = 3.769, SD = 0.765). Those with prior motion control experience also reported the sound to be more responsive (M = 3.846, SD = 0.909) than those without (M = 3.154, SD = 0.834). They also reported the audio-visual experience felt more synchronised (M = 3.577, SD = 1.065) than those without (M = 3.462, SD = 1.104).

15 out of 26 participants (57.69%) had experience playing a musical instrument. Of these 15, 10 (75%) had played for over seven years. Only one of the remaining five participants had played for less than a year, with the remaining participants playing between one and seven years. Those with musical experience reported that it felt less like a musical experience (M = 2.533, SD = 1.060) than those without musical experience (M = 3.182, SD = 1.052). A reason for this could be those with musical experience would have a higher demand for what a 'musical experience' would contain. They would likely expect more creative agency or range of expressivity for their interaction in order for them to consider it musical. Contrary to this, participants without experience may have fewer preconceived notions of what a musical experience should contain and reduce it down to a more passive experience. Those with more musical experience also reported that missing a target did not break their focus (M = 3.267, SD = 1.060), compared to those without experience (M = 3.5, SD = 1.102). This is not surprising, as musicians are trained to remain unfazed by making a mistake in performance. 18 out of 26 participants (69.23%) had experience using music production software, and these participants also reported it felt less like a musical experience (M = 2.311) compared to those

without music production software experience (M = 3.813). Those with music production software experience were also much more likely to make use of the visual cues (flashing playheads) to guide their input (M = 2.361) compared to those without (M = 3.75). Across the full 31-question IEQ, these players with music production software experience reported a lower sense of immersion (M = 3.203, SD = 0.343) than those without (M = 3.611, SD = 0.472). This supports the notion that unless a musical challenge meets the expectation of a player with experience, it will impact their experience negatively. Despite both groups finding the overall game similarly challenging (with experience M = 3.778, without M = 3.938). However, it is worth considering these groups are not equally weighted.

Evaluation

The results between the two distinct groups across all sections of the survey indicate the playing experience has changed between the two versions of the game, as shown in the results of the IEQ in Table 32 and 33 on page 151. However, the results do not uniformly show one version of the game as having a clearly improved playing experience and sense of immersion across each section of the survey. While the analysis of the combined sections shows Game B as having a slightly higher reported value of playing experience (M = 3.239) than Game A (M = 3.242), the difference is small. Yet in section one of the survey, Game A was reported to have a greater playing experience (M = 3.418) compared to Game B (M = 3.231) by a higher margin, indicating there was a larger difference relating to the playing experience. Contrary to the results regarding the playing experience, the second section of the survey results were inverse, with players reporting a greater sense of immersion in Game B (M = 3.245) compared to Game A (M = 3.188). This was further compounded by the findings of the complete 31-question IEQ, where Game B (M = 3.343) was found to be a more immersive experience than Game A (M = 3.115), although these results are significantly closer than the results of the optimised second section. This elucidates a complex relationship between the playing experience and sense of immersion, where certain demands from the playing experience, along with prior experience of the player, impact immersion.

In the first section of the survey regarding playing experience, players found Game A to be a more musical experience (M = 2.923) compared to Game B (M = 2.692). This is despite 57.69% of the participants with musical experience reporting that it felt less like a musical experience (M = 2.533, SD = 1.060) than those without musical experience (M = 3.182, SD = 1.052). The interaction method (pushing one of two triggers) was the same for both versions of the game, with the only change in sound design as an outcome. One can attribute this greater sense of musicality to the triggered audible event being a note relative to the non-diegetic soundtrack in Game A as opposed to a sound effect. We can also attribute the interaction from the player being more consequential to the musical experience, as their timing or synchronisation of shooting becomes musically performative. The same quality of the interaction is also responsible for why players found Game A (M = 3.7308) to have more responsive sound than Game B (M = 3.308). By virtue of the player's interaction being musically performative, it also led to players focusing on synchronising their inputs to a greater extent in Game A (M = 3.538) than in Game B (M = 3.077). One can argue these results demonstrate a more engaging playing experience in Game A; the players are focusing more of their attention on the music experience, as well as directly on the gaming challenge. This is despite the players having a closer audio-visual relationship in Game B, where there is a direct one-to-one relationship between the music backing track and the 3D graphics that make up the target. These results show there is more value to the playing experience in encompassing musical interaction than by minimising it.

However, this additional context of musical interaction potentially comes at a cost to the playing experience. Making players responsible for a more consequential sound design feature led to a lower sense of synchronisation between the audio-visual elements in Game A (M = 3.115) than in Game B (M = 3.923). This result is unsurprising given all players were new and inexperienced to the game. Subsequently, they are unlikely to perform as accurately as required to remain synchronised with the predetermined musical arrangement. If players were given more time to practice, these results could change. We can also attribute this inexperience, along with added musical interaction as discussed above, as a reason players reported that missing a target/block and taking damage broke their focus in Game A (M = 4.039) considerably more than in Game B (M = 2.692). This argument is also supported by cross-referencing the demographic characteristic data. As mentioned in this chapter earlier, those with more musical experience were less likely to break their focus when taking damage compared to those without experience (with experience M = 3.267), (without experience M = 3.5). This tells us that even comparable experience is useful for synchronised performance in our game.

Another consequence of increasing the musical interaction in this manner is it impacted the player's immersion negatively. This was shown in both our results for section two (Game A M = 3.188, Game B M = 3.245) and in the complete IEQ, which showed a greater level of immersion in Game B overall (Game A M = 3.315, Game B M=3.343) and specifically in three of the five factors: 'Cognitive Involvement', 'Emotional Involvement', and 'Control'. One can argue the drop in immersion for Game A in 'Cognitive Involvement' (Game A M = 3.449, Game B M = 3.58) is a result of the increased perceptual and cognitive overhead of simultaneously parsing 3D targets to shoot, along with the cognitive processing of music (that is moving in and out of sync depending on how well the player is performing). When players perform poorly (and as they were all inexperienced at the game, it is fair to say many of them did), it increases the cognitive demand beyond a limit suitable for immersion. This could be mitigated by quantising player input or using the player input to confirm future arranged elements, as opposed to triggering them live in sequence. This is also supported by the results in section one, where players reported the interaction was more intuitive when the interaction triggered a sound effect in Game B (M = 4.192) rather than music in Game A (M = 3.731).

In the factor of 'Emotional Involvement' (Game A M = 2.712, Game B M = 2.795), one could infer the increase due to the affective nature of the complete in-tune and in-time backing music of Game B against the inconsistent player-assisted backing music of Game A. It is also salient to remember Emotional Involvement also includes negative emotions, not just positive ones. This is reflected in the desire to play again when asked; players were less likely to play Game A (M = 3.38) than Game B (M = 3.615). However, in section two, players reported they felt more disappointed when the game finished in Game A (M=2.932) than in Game B (M = 2.577). We could attribute this to the novelty of the game design in the A version, as opposed to the typical rhythm/music game design of Game B. Further results support this interpretation as players also reported they were more motivated in Game A (M = 3.230) than in Game B (M = 3) and expressed a higher desire to give up during Game B (M = 3.577) than in Game A (M = 3). However, players also reported they felt like they were making more progress towards the end of the game in Game B (M = 3.461) as opposed to Game A (M = 2.6358). Considering player performance would likely improve upon further play, these results could change over a longer duration.

The immersion factor of ‘Control’ was remarkably close between the two versions of the game (Game A M = 3.077, Game B M = 3.131). Players felt like they were moving through the game according to their own will more in Game A (M = 2.308) than in Game B (M = 2). As player movement in both games is fixed, and the targets are presented in the same position and order, one can infer this difference only occurs due to the change in sound design. By changing the outcome of the player’s interaction from diegetic sound effects to musical accompaniment and assisted arrangement, players felt like they had more agency in the game. However, despite sharing the exact same control scheme, as reported above when discussing Cognitive Involvement, the potential added complexity of processing the Game A challenge impacted the playing experience. In relation to the control factor, players felt more likely to become so involved that they were unaware they were even using controls in Game B (M = 2.462) than in Game A (M = 2.192). In addition, despite featuring the same graphics and imagery for both games (albeit with the arrangement sequenced with a starting offset to allow for the synchronised playhead to be visible to the player in Game B), players reported they enjoyed the graphic and imagery more in Game B (M = 3.192) than in Game A (M = 2.885). This could be considered an example of transference across sensory modalities in perception.

In the remaining two factors of ‘Real World Dissociation’ and ‘Challenge’, Game A was found to be more immersive than Game B. As discussed in the previous study on page 122, it is not surprising that both games were shown to be highly immersive in the factor of ‘Real World Dissociation’ (Game A M = 3.748, Game B M = 3.736) as both games are played in VR, in a headset that covers your eyes. Participants were also asked to turn the volume up loud so they could hear the music. This initial starting position of sensory deprivation will lend itself to immersive experiences, as once the game starts, there will likely be fewer sensory distractions for the player. This is coupled with a strong desire to win (Game A M = 3.769, Game B M = 3.846), and one can expect that players would not be interested in their surroundings while playing the game. We can relate the final factor of 'Challenge' to the same additional musical 'challenge' of the player's requirements of performance (with greater consequence for their actions) in Game A (M = 3.462), compared to Game B (M = 3.231). While this is a negative for the immersion factor of Cognitive Involvement, the same task is responsible for a greater sense of immersion in the factor of Challenge. However, the net difference between the two factors is a positive of 0.112 in favour of Game A, meaning when considering both factors of Cognitive Involvement and Challenge, Game A is a more immersive experience. This is also further supported in our discussion of the playing experience above. As discussed in the data analysis of the demographic characteristics, players with more experience with motion controls and those with music production software may already be familiar with musical concepts found within this game reported a lower sense of immersion. Likewise, as discussed in the demographic analysis, those with musical experience were less likely to find it engaging and immersive. An argument could be made that this is due to the player's expectation of their own performance rather than a lack of agency in the interaction. However, one possible solution could be increasing the musical difficulty via a sliding scale. Those with no musical experience maintain the synchronised sequence matching until they improve; those with lots of experience could have agency increased; rather than sequence matching, they could be asked to shoot the target blocks with different notes/buttons.

This has important consequences for game development regarding sound design, along with specific implications for music game challenge design and level design. If it proves worthwhile to increase the impact of player interaction with regard to sound design, then further work should incorporate and document this in greater detail. Any application of this approach has the potential to change the game design process significantly. Designers would need to consider in greater detail how responsive their interactions were, with both early development testing and later Q&A testing checking to see how performant the interaction was. Further user testing may also be required, as each player's grounding of sonic expectation will be different, as per their own cultural experiences (Klarlund et al., 2023) (Huron, 2008). Any additional programmatic or procedural levels of behaviour will add further complexity. All of which would increase the project lead time for completion.

Conclusion

By changing the sound design of the player interaction, we can impact the playing experience and the player's sense of immersion significantly in many areas. However, this impact is not uniform across the five factors of immersion, as presented in the IEQ. By attaching a musical non-diegetic outcome, as opposed to a diegetic sound effect, to player interaction, we noticeably impacted Cognitive Involvement negatively and the Challenge factory positively. To a lesser extent, we also impact Control and the Emotional Involvement of the player. We can also conclude immersion and the playing experience are interconnected; players reported a more musical and engaging playing experience but also found it to be less immersive, especially when they had a more traditional musical experience. To what extent is this dependent on the novelty of the gaming challenge? If players perform better over time, will this lead to a greater sense of immersion? Do the factors of Cognitive Involvement, Control, and Challenge scale when increasing the complexity of musical interaction? By expanding this research and answering these questions, we will gain a better understanding of the role of game design, sonic interaction, and immersion.

Discussion

Research Questions and Objectives

The first aim of this research was to further the understanding of how the playing experience in games is impacted by sonic expression. The motivation was to improve knowledge of sound design in games and provide inspiration for the further use of interactive sound and novel game design. By developing and testing three experimental games, this research has provided valuable data that can be used to address this aim and has identified factors of sonic expression that influence the playing experience. The second aim was to provide insight into the role of sonic expression in player immersion. By utilising interactive sound throughout all three games and applying a variety of different approaches to game design coupled with different sound design techniques, this research provides ample data to elucidate the impact sonic expression can have on player immersion. By analysing and evaluating the impact sonic expression has on the player's sense of immersion, we can inform game design practice. This should encourage the future development of games that deliver an enhanced playing experience and with a greater sense of immersion for the player. The third aim was to ascertain the relationship between sonic expression and gaming challenges in the context of game design. By using different game design paradigms of a puzzle game, a 3D platform game, and a rhythm game, this research provides an example of how sonic expression impacts the gaming challenge across design strategies. This provides discrete insights for music game design and a broader understanding of novel methods for game challenge design. The outcome of these findings can be used to develop a conceptual framework to aid game designers in incorporating the use of sonic expression within game design practice.

Summary of Findings

The Preliminary Study was used to explore the player's experience and sense of immersion when their sonic expression was used as a game mechanic in a puzzle game. It was also used to determine the appropriate platform for the game development within this research. The study suggested there was a potential for music games to be better experienced in VR and found that VR lent itself to a comparatively more immersive experience than that of a traditional screen (M = -0.3917, SE = 0.370, t(11) = 3.672, $p < 0.05$, with an effect size d = -1.075). This was supported by findings from scoring the 31 IEQ (VR M = 3.658, Screen M = 3.306). The study also highlighted potential difficulties regarding the design of music games that rely on the player's sonic expression, specifically as the primary in-game challenge in the context of a puzzle game. Players reported that the screen-based game was more challenging (M = 4.333) compared to the VR version (M = 3.167). They also felt more in control in the VR version (M = 3.833) than the screen-based version (M = 2.083). Despite these difficulties, players were engaged and wanted to win the game equally across both versions of the game (VR M = 3.5, Screen M = 3.5). Players also reported the controls were more intuitive in the VR version (screen M=3.35, VR M=4 Q6 section one). This compounded the findings relating to immersion and directed the game development towards VR platforms for the following studies.

The second study was used to ascertain the role of player agency and sonic expression, and the subsequent impact it has on the playing experience when used as a gaming mechanic for a 3D platform game. It explored to what extent sonic expression could be leveraged in game design, along with the corresponding impact this has on the sense of immersion. The outcome helped establish a relationship between the player's musical-expressive decisions, as opposed to their desire to complete an in-game challenge. The study found that, by giving the players greater agency in sonic expression, we can significantly increase the player's sense of immersion. This was reflected in the findings of the second section of the post-survey (Game A section two M = 3.28, Game B section two M = 2.830) and in the scoring of the IEQ (Game A M = 3.15, Game B M = 2.85). This was also reflected in the playing experience, with a substantial impact on player motivation (Game A M = 3.5, Game B M = 2.767) and enjoyment (Game A (M = 3.367, Game B M = 2.767). Players also expressed a much higher desire to replay the game that had increased agency (Game A M = 3.4, Game B M = 2.5). The study also highlighted that by increasing agency, players found the game more challenging (Game A M = 3.77, Game B M = 2.9). This was also reflected in their sense of immersion within the immersion factor of the challenge of the IEQ (Game A M = 3.333, Game B M = 2.625).

The third study uncovers the impact on the playing experience when player interaction triggers either diegetic or non-diegetic sound in response to synchronised audio-visual cues within a rhythm-shooter game. The study explored the impact of sound design on the playing experience and the player's sense of immersion. Additionally, the study explored the role of synchronised audio-visual cues in facilitating the playing experience. The study found that changing the sound design of player interaction impacted the player's sense of immersion. However, the impact was not uniform. The use of non-diegetic musical sound for player interaction negatively impacted the IEQ immersion factor of cognitive involvement (Game A M = 3.449, Game B M = 3.568) and control (Game A M = 3.077, Game B M = 3.131) but impacted the factors of challenge positively (Game A M = 3.462. Game B M = 3.231). The study found that the immersion of individual players was also impacted by their prior musical experience. Players with music production software experience reported a lower sense of immersion (M = 3.203, SD = 0.343) than those without (M = 3.611, SD = 0.472). Despite reporting a lower sense of immersion, players reported a more engaging experience when their interaction triggered non-diegetic musical sound (M = 0.187, SE = 0.677, $t(25) = 1.407$, $p < 0.05$, with an effect size $d = 0.677$). Players also reported that when their actions generated non-diegetic sound to form part of the backing music, they focused more on synchronising their playing inputs. Additionally, when asked, 'What sensory feedback was the most useful in guiding your input?' 50% of players reported they used auditory cues of the backing music to guide input, as opposed to visual cues. Whereas in game B, only 23.08% reported using the auditory cues, despite having the complete backing track only in game B.

To What Extent Does Sonic Expression Impact the Playing Experience?

The findings of the three studies show that sonic expression had a substantial impact on the playing experience and the player's sense of immersion. In relation to the research question of 'To what extent does sonic expression impact the playing experience?' the Preliminary Study found that when integrated into gameplay, sonic expression broadly contributed positively to the playing experience. The findings highlight that using sonic expression as a gaming mechanic has the potential to increase the difficulty of the in-gaming challenge. As discussed in the literature review on page 35, corporeal music engagement is confined to small, intended actions, the corresponding interaction, the consequential descriptors (Wanderley & Orio, 2002), and the accompanying in-game challenge. If sonic expression as a gaming mechanic does increase the difficulty of the in-game challenge, then this highlights an interesting observation: is there a possible conflict between auditory perception and cognition and the cognition involvement for spatial reasoning as contextualised in the game challenge? When the embodied experience relates to the corporeal music engagement and the corresponding physical action that has a corresponding game action that does not deviate, why would there be a conflict? This could be explained by the player finding the mapping of the device (Fels et al., 2002), i.e., the technical affordance (Gibson, 1979) (Gaver, 1991) of the in-game sound interaction, to be insufficient or incomplete relative to their expectation. (Tanaka, 2010) (Poepel, 2005). This also relates to the specific sounds used within the game (Pichlmair & Kayali, 2007) (Kurtz, 1998). To confirm this theory, one would need to de-couple the interactive gaming challenge of organising cubes in specific formations from the current musical challenge of developing a coherent musical arrangement. This would be possible but would require additional visual cues to frame the challenge, along with a different motivation or prompt.

The findings show that players found the controls more intuitive in the VR version (VR M = 4, Screen M = 3.35). This demonstrates that it was not solely the interface or in-game interaction design that was responsible for the gaming challenge being difficult. We can argue that the VR version is more intuitive as the quality of device mapping (Fels et al., 2002) is comparatively better than that of the screen-based solution. This was also reflected in the IEQ scoring data for the immersion factor of control (VR M = 3.7, Screen M = 3.05). In the data analysis by this researcher in section two of the survey and the IEQ scoring, the VR version of the game was found to be conducive to player immersion and positively impacted the playing experience. Players felt more immersed as they did not need to focus attention on their input. This is not surprising, given the similarities between selective attention and immersion, as described by Jennett, where 'manipulating features that are known to influence game immersion, such as a person's sense of progression, one is able to affect how much people attend to other aspects of their environment' (Jennett 2010). The evidence further supports the assertion that the method of interaction was not responsible for the apparent difficulty of the task with the results of the Preliminary Study regarding immersion, as interaction must meet expectations for immersion to occur (McMahan, 2003). The overall findings relating to immersion can also framed within Cairn's hierarchy of immersion, where immersion corresponds to the player's sense of engagement and involvement. Players move from a lower level of immersion, investing time and effort towards a higher level of immersion beyond dedicated attention into presence (Cairns et al., 2014). 'Total immersion was seen as the idea of complete involvement with the game where nothing else matters and the player feels "in the game."' (Cairns et al., 2014). As discussed in the study evaluation on page 92, it is not surprising players reported high levels of immersion in the VR version, as the headset, by design, blocks out your surroundings, along with potential distractions around you. This is compounded in a music game, where game audio loudly plays directly from the headset, masking any background noise. This was reflected in the IEQ scoring for the immersion factor of real-world dissociation (VR M = 3.917, Screen M = 3.155). Subsequently, it was concluded that VR provided a more suitable platform compared to a screen-based medium for developing games that harness sonic expression as a game mechanic, as the medium was conducive to player immersion and positively impacted the playing experience. This is noteworthy for future music game developers, as they will find better player engagement for their games when using the most appropriate platform for expressive sound.

In the second study, we found that increasing player agency in sonic expression significantly increased the overall playing experience and the player's sense of immersion. Players were more motivated when they had greater depth of sonic expression and reported higher levels of enjoyment. We can attribute this to two different factors. The first factor is that this game has a comparatively more clearly designed game challenge than the puzzle challenge. The puzzle game challenge was to 'score points by building a coherent rhythm' (page 228). This is heavily contextualised in music, relying on the player's understanding of 'a coherent rhythm'. As music is encultured (Klarlund et al., 2023) (Huron, 2008), this may cause issues. Additionally, while players were able to experiment and learn the game through the tutorial tasks, the relationship between their sonic expression and the points earned is primarily reliant on their music comprehension. Although the game featured a 3D graphical representation of their arrangement, it did not actively guide the players towards 'good' future placements and was merely representational of their previous interaction. This is a drastic departure from existing rhythm games such as Guitar Hero (Harmonix, 2005), Parappa the Rapper (NanaOn-Sha, 1996), or Vectronom (Ludopium, 2019). Subsequently, we can say the in-game challenge was more musical than the traditional and expected music-game challenge. Without a firm understanding of music, this could cause players to struggle. Additionally, players with musical experience who also regularly played games will likely have an expectation for the experience to be game-like, as opposed to musical. This would also impact their playing experience in a similar manner. The in-game challenge in the 3D platformer prototype addresses this issue. By changing the challenge so it is closer to a traditional challenge found in a platform game - navigating from a defined start point to a defined endpoint, the player ends up with a better understanding of the challenge set, resulting in higher levels of enjoyment. Importantly, by having a clearer game challenge for the platform game, when the players were given increased agency and a greater depth of sonic expression, their enjoyment increased (Game A (M = 3.367, Game B M = 2.767).

The second factor responsible for increasing the overall playing experience is the resultant affordance (Gibson, 1979) (Gaver, 1991) of the game as the player agency is increased, correspondingly increasing their range of sonic expression (Tanaka, 2010) (Poepel, 2005). By giving players a gaming mechanism that offers more opportunities for musical experimentation, players were more motivated to engage with the game (Game A M = 3.5, Game B M = 2.767). This is also reflected in the player's sense of immersion (Game A section two M = 3.28, Game B section two M = 2.830). It is not unexpected that we see an increase in motivation reflect a higher level of immersion as 'a willingness to concentrate' (Brown & Cairns, 2004) is vital. When players were given a broader range of sonic expressions, they reported a higher level of immersion for the IEQ immersion factor of control (Game A M = 3.386, Game B M = 3.02). As covered in the literature review on page 47, the SCI model (Ermi & Mäyrä, 2005) of immersion summarises immersion to be a sum of varying levels of three immersion components: sensory, challenge-based, and imaginative (Ermi & Mäyrä, 2005). Sensory immersion pertains to the audio-visual experience of the game, challenge-based relates to the motor or spatial reasoning utilised, and imaginative relatives to narrative or affection. Increasing player agency for sonic expression impacts each of these three components, as it gives the player a broader range of audio-visual experience, a higher dimensionality of spatial reasoning (as they have a larger number of decisions to make), and they become responsible for the musical (narrative) path they take. This can be considered ‘dedicating attention’ (Jennett et al., 2008) (Grodal, 2003) (Klimmt, 2003), which would reach the level of engrossment (Brown & Cairns, 2004) in Cairn’s hierarchy. As discussed in the literature review on page 45, an optimised SCI model can be considered total immersion in Cairn’s hierarchy. While increasing player agency positively impacted the playing experience and immersion, it also increased the perceived difficulty of the game (Game A M = 3.77, Game B M = 2.9). This reaffirms the importance of the in-game challenge as discussed above and reaffirms the complexity of using sonic expression as part of an in-game challenge found in the Preliminary Study.

In the third study, we found that by changing the quality of sound triggered by the player’s interaction, we impacted the playing experience and the player’s sense of immersion. This highlights the importance of sound in the playing experience (Tan, 2014). By A/B testing a diegetic sound effect against a non-diegetic musical response, the results show the impact on playing experience and immersion when sound design deviates from traditional approaches as they are defined in film theory and practice (Chion & Gorbman, 1994). While there are existing sound-design frameworks such as the IEZA (Huiberts & van Tol, 2008) that add depth to this approach, outlining specific domains of ‘Effect, Zone, Interface, and Affect’ that relate to the player’s activity, the approach is not a radical departure from sound-design in film. The IEZA is primarily concerned with optimising gameplay (as expressed by Cunningham regarding supporting playability (Cunningham, 2006)), as well as to ‘dynamise gameplay’, meaning to make the gameplay experience more engaging (Huiberts, 2010). This researcher would argue that to follow the IEZA framework, the game design of the prototype in the third study would require an additional context for why shooting the targets triggers music. By not providing this context, we are directly challenging the validity of the framework and the value of treating sound design as two conceptual independent domains (non-diegetic or diegetic). The findings reflect that the effects of the change are nuanced. The results demonstrate that the type of sound expressed by the player has a distinct impact on differing facets of immersion, along with the overall playing experience. This outcome strengthens the findings from the existing research on the influence of sound on playability by Yamada (Yamada et al., 2001) (Yamada, 2002) and Lipscomb & Zehnder’s study on the impact of a musical score on the gaming experience (Lipscomb & Zehnder, 2005), as it highlights the impact specifically on immersion. Additionally, it was found that triggering non-diegetic musical events, as opposed to a diegetic sound effect, led to a more engaging experience despite a marginally less immersive experience overall. This is a valuable insight for future game design, as it highlights the worth of deviating from existing sound-design conventions.

Does Increasing the Depth of Sonic Expression Produce a Measurable Change in Immersion?

In relation to the research question, 'Does increasing the depth of sonic expression produce a measurable change in player immersion? while the Preliminary Study did not explicitly investigate the depth of sonic expression, as both versions of the game delivered the same range of expression to the player, it did indicate that VR offers the possibility of a more immersive experience for games that feature sonic expression (Screen vs VR $M = -0.3917$, $SE = 0.370$, $t(11) = 3.672$, $p < 0.05$, with an effect size $d = -1.075$). This is not due to a lack of expressivity, as both versions of the game shared the same design and were effective at conveying meaning or feeling (Dobrian & Koppelman, 2006). Instead, this highlights the importance of the embodied interaction in the role of immersion. As discussed above, when addressing the research question pertaining to the impact of sonic expression on the playing experience, interactions must meet expectations for immersion to occur (McMahan, 2003). The findings support the notion that the technological affordance (Gibson, 1979) (Gaver, 1991) provided by a VR platform has a higher ceiling for immersion than that of screen-based platforms when delivering sonically expressive interactions.

Using the SCI model (Ermi & Mäyrä, 2005) for immersion, we can frame the VR version as scoring higher in sensory, challenge and imaginative components. This is due to the nature of wearing a VR headset, where the player's sensory experience is dominated by the game environment. It is also due to the size of the 3D graphics inside the game environment. Players are required to make a much larger physical movement to achieve their intended actions. Given the importance of gesture on broad interaction within our corporeal music engagement (Leman, 2007) (Wanderley & Orio, 2002), one can argue the increase in the player's sense of immersion and the corresponding interaction method between the two game versions is tantamount to a range of expression. Subsequently, we have produced a measure change in immersion due to an increase in the depth of sonic expression (an increase in the sensory and motor skills required for interaction).

The second study found that increasing player agency and broadening the impact of the player's expression led to significant measurable change in player immersion (Game A section two M = 3.28, Game B section two M = 2.830). Players exhibited heightened immersion across immersion factors of the IEQ (Game A M = 3.15, Game B M = 2.85). These findings clearly show when the technological affordance (Gibson, 1979) (Gaver, 1991) offered by the game design changes, immersion changes as well. There are two factors that influence this. The first factor involves the increased device mapping (Fels et al., 2002) and the corresponding increase in functionality that affords the player to situation their expression (Tanaka, 2010) (Poepel, 2005). This also relies on the player's ability to navigate the space between interaction and sound (Leman, 2007), as a lack of comprehension to do so would negate the impact of the increased technological affordance (Gibson, 1979) (Gaver, 1991). This highlights the importance of 'the transparency of device mapping' (Fels et al., 2002) because a device that is difficult to understand and use will likely fail at delivering an expressive experience.

The second factor is how the depth of sonic expression offered to the player relates to the in-game challenge design, as this also impacts player immersion too. As discussed in the literature review on page 37, existing rhythm games 'offer little freedom of expression apart from the prerogative to perform while playing. They strictly force rules on the player on how she has to react to a specific stimulus displayed on screen or communicated by sound.' (Pichlmair & Kayali, 2007). In rhythm games such as Guitar-Hero (Harmonix, 2005), this is because the game challenge is often a series of strict sequence matching, where any deviation of input is punished. In rhythm games such as Vib-Ribbon (NanaOn-Sha, 1999) or Audiosurf (Fitterer, 2008), the audio file is used to procedurally generate the level to play. Once generated, the games follow a similar pattern of performance to Guitar-Hero; the player follows strict rules. By deliberately designing the game challenge to be open, offering a much greater depth of expression, the games in this thesis allowed players to express themselves musically and create their own path to the endpoint. This was achieved by using procedural content generation (Hendrikx et al., 2012) (Risi et al., 2014), generating the platforms (and enemies) based on the player's sonic expression. This approach was inspired by challenge modelling for procedural level creation covered by both Sorenson (Sorenson et al., 2011) and Smith (Smith et al., 2011). Although the implementation of procedural content generation was considerably simpler in this research, the approach still yielded interesting results. It also presented them with a novel playing experience every play through if they changed their expression. This likely forced the players to be more attentive, resulting in the higher levels of immersion we see clearly in the results (Jennett et al., 2008) (Grodal, 2003) (Klimmt, 2003).

Although the third study did not explicitly focus on increasing the range of sonic expression within the game challenge, it did focus on measuring changes in immersion when two differing sound-design approaches were undertaken (triggering non-diegetic musical events or diegetic sound effect events). We can consider these changes to be broadening the depth of sonic expression, as the measurable change directly corresponds to the quality of the sound. These changes highlight meaningful characteristics for experimentation within sound design and help further understanding of the role of sound design in immersion. This is especially important as, despite the sound being an important component of the overall game aesthetics and affective perception (Lennart et al., 2011), there is little theoretical support for game designers (Alves & Roque, 2011) (Collins, 2015). From the results, we can determine that the player's sense of immersion was impacted negatively by their prior musical experience, especially when players had previous experience with music production software (with experience M = 3.203, without experience M = 2.311). This same group also reported that it felt less like a musical experience (M = 2.311) compared to those without music production software experience (M = 3.813). This negative impact on immersion is likely explained by the differences in expectations between the experienced player and the game environment. Players with music software experience would also likely find the interaction trivial in comparison to music production software. This is supported by two of McMahan's three conditions for immersion to occur: 'the user's expectations of the game or environment must match the environment's conventions fairly closely; the user's actions must have a non-trivial impact on the environment.' (McMahan, 2003).

The two different sound-design approaches also resulted in one of the game versions potentially having a different arrangement of the bassline within the overall game music. Players in Game A would need to synchronise their inputs in the appropriate place for each target to achieve the same arrangement as Game B. While this offers the opportunity for the player to be more expressive in their rhythmic input and 'evoke a variety of rhythmic qualities, accents, or emotional moods by playing notes slightly late or early relative to a theoretical metric time point.' (Iyer, 2002). The player only had a short window of time to make this interaction, with a high number of targets with which to interact with that shared similar behaviour (pitch and duration changed for each target, but many shared the same pitch and duration). Subsequently, we can claim that the technological affordance (Gibson, 1979) (Gaver, 1991) was not sufficient for situating the player's input to the sound and corresponding game response (Tanaka, 2010) (Poepel, 2005), resulting in the player's lack of immersion due to their 'non-trivial impact on the environment' (McMahan, 2003).

How Does Sonic Expression Impact Game Challenge, and Does This Affect Immersion?

In relation to this research question, the Preliminary Study highlighted potential difficulties when designing games that rely on the player's sonic expression as the primary in-game challenge. Players were engaged equally across both versions (VR M = 3.5, Screen M = 3.5) and wanted to beat the in-game challenge. However, players found the Screen-based version considerably more challenging (Screen M = 4.333) compared to the VR version (M = 3.167). The high level of perceived difficulty we observe in the results is likely due to the abstract nature of the game challenge. As discussed in this chapter above, the challenge was heavily contextualised in music, relying on the player's understanding of rhythm. We could frame this differently regarding the game challenge, stating that a puzzle game was not the most appropriate genre of game for this mapped technological affordance (Gibson, 1979) (Gaver, 1991) that was offered in the sound and corresponding challenge function of the game (Tanaka, 2010) (Poepel, 2005). Existing research by Yamada and Lipscomb & Zehnder concludes there is a complex relationship between affection playability between audio and visual components (Yamada et al., 2001) (Yamada, 2002) (Lipscomb & Zehnder, 2005). Lipscomb & Zehnder's research found responses that include transference between audio and visual components. If transference between sensory modalities is possible, it follows that interference from other sensory modalities is also plausible.

Another potential explanation for why using sonic expression as a primary gaming mechanic for a puzzle game is not the most appropriate choice for an enjoyable, intuitive challenge is that sonic expression is confined to temporality. Whereas puzzles often require meditation, many observations, and sometimes require to be viewed from multiple perspectives to solve. This leads to a break in expectation, which in turn breaks immersion (McMahan, 2003). The self-imposed internal challenge of 'the act of communicating' (Fels et al., 2002) that is derived through sonic expression is incongruent with typical gaming puzzles. As discussed in above for the previous research question, the measurable effects of sonic expression in relation to the gaming challenge affect immersion. Despite the results regarding the game challenge, players still reported decent levels of engagement and motivation to continue playing. This suggests there is a complex relationship between sonic expression and in-game challenge.

In the second study, increasing player agency in sonic expression also led to an increased perception of the game's challenge. This relationship between player agency challenge and immersion demonstrated that the role of sonic expression in shaping the game's challenges has significant implications for the player's sense of immersion. Although players found the game more challenging when given increased levels of agency for their sonic expression (Game A M = 3.77, Game B M = 2.9), players also found the experience more immersive (IEQ Challenge factor Game A M = 3.333, Game B M = 2.625). This supports the findings from the previous study, along with the existing research regarding the impact of game genre as a factor in how sound can influence the player's experience or ability to play the game (Yamada et al., 2001) (Yamada, 2002) as it demonstrates that not all gaming paradigms respond to the use of sound in the same way. An argument that counters this conclusion (although not music-exclusive) would be the difference in the levels of sophistication between the 3D graphics in each game. The 3D platform game has a more complex and visually imposing game environment. The graphics are larger, with a clear distinction between the visual representation of the different sound elements available to the player. Importantly, they are comprised of a combination of fixed and audio-visually synchronised moving 3D graphics. The puzzle game in the Preliminary Study had comparatively smaller graphics, did not use colour to differentiate types of sound, and synchronised the movement of all the arranged 3D graphics. As discussed above, this was not conducive to making the gaming challenge clear and accessible and impacted immersion accordingly.

The third study directly examined the relationship between sound design (inclusive of sonic expression), challenge, and immersion. It shows that sound design choices impact the perception of challenge and, in turn, influence immersion. Players reported that when their actions generated non-diegetic sound to form part of the backing music, they focused more on synchronising their playing inputs. This is perhaps counter-intuitive, as one would expect players would be more likely to synchronise their input to a backing track that contained all arranged musical elements, as opposed to the backing track that was missing the player's input. This desire to synchronise input in this instance could be tantamount to the gestalt principle of closure (Chang & Nesbitt, 2006) and illustrative of the player's motivation to bridge the gaps as identified through their imagination (Ermi & Mäyrä, 2005). As discussed, when addressing the second research question in this chapter, players had a more engaging playing experience when triggering non-diegetic sound.

However, this was not reflected uniformly in their sense of immersion. As covered in the findings, prior relative experience impacts the player's immersion negatively, as players determined that their action had a 'trivial impact on the environment' (McMahan, 2003). This was also the case regarding how much of a musical experience the games felt. Overall, players found Game A to be a more musical experience (M = 2.923) compared to Game B (M = 2.692). This is despite 57.69% of the participants with musical experience reporting that it felt less like a musical experience (M = 2.533, SD = 1.060) than those without musical experience (M = 3.182, SD = 1.052). This shows the importance of the novelty of the experience, especially when concerning play. Ermi and Mäyrä relate McMahan's 'non-trivial impact on the environment' to Salen and Zimmerman's (Salen & Zimmerman, 2003) definition of meaningful play, stating that it 'occurs when the relationships between actions and outcomes are both discernible and integrated. Discernibility means letting the player know what happens when they take action, and integration means tying those actions and outcomes into the larger context of the game.' (Ermi & Mäyrä, 2005). Consequently, we can argue that players with prior experience found both versions of the game to have little discernible action and were not satisfied with the level of integration. Comparatively, players without prior experience found that the non-diegetic version of the game had a discernible action and were satisfied with the level of integration (if not finding their attention more frequently disrupted when missing a target (M = 3.5) compared to those with experience (M = 3.267)) compared to the diegetic version of the game. These conclusions share similarities of demands that could be related back to design theory and the technological affordance of devices for musical expression (Wanderley & Orio, 2002) (Tanaka, 2010).

A Conceptual Framework for the Use of Sonic Expression in Game Design

The outcome of this research can be used as a basis for developing a conceptual framework for the use of sonic expression in game design. Through interpreting the results, we can elucidate that when using sonic expression, the expectation of the interaction must be met (McMahan, 2003) within the context of the game environment and the player's auditory cognition. Their sonic expression must make a discernible impact on the game environment, and there must be sufficient technological affordance (Gibson, 1979) (Gaver, 1991) for the player to feel that this is in their control. As an issue to overcome, this primarily relates to the design of in-game challenges, along with the parallel sound-design methods used in the game. When incorporating sonic expression within a gaming mechanic, it is beneficial to allow the player to openly explore the use of their expression at their own pace. It is also important to provide a clearly defined challenge that can be easily conceptualised outside of sound perception and cognition. The mapping (Fels et al., 2002) of affordance and the affordance itself must not be too limited. Players should be able to intuitively identify appropriate actions and understand how their actions navigate the gaming challenge. Appropriate game design methods can be harnessed to accomplish solve these issues.

The framework can be divided into the following four connected domains:

- Demand: the necessary sonic expression used to complete the in-game challenge should be reasonable, and challenges should not be overly performative in the sound domain.
- Inclusivity: in-game challenges should make use of sonic expression to navigate the challenge, but sonic expression should not be a crux; sound design choices should frame the challenge.

- Versatility: the method to navigate through the sound-challenge space should be open. It should allow the players to solve challenges in a variety of different ways.
- Engagement: sonic expression should be continuous, and relative actions clear. The sound design approach should not disrupt or interfere with the challenge.

An overview of how the three studies feed into the framework is shown in the diagram below.

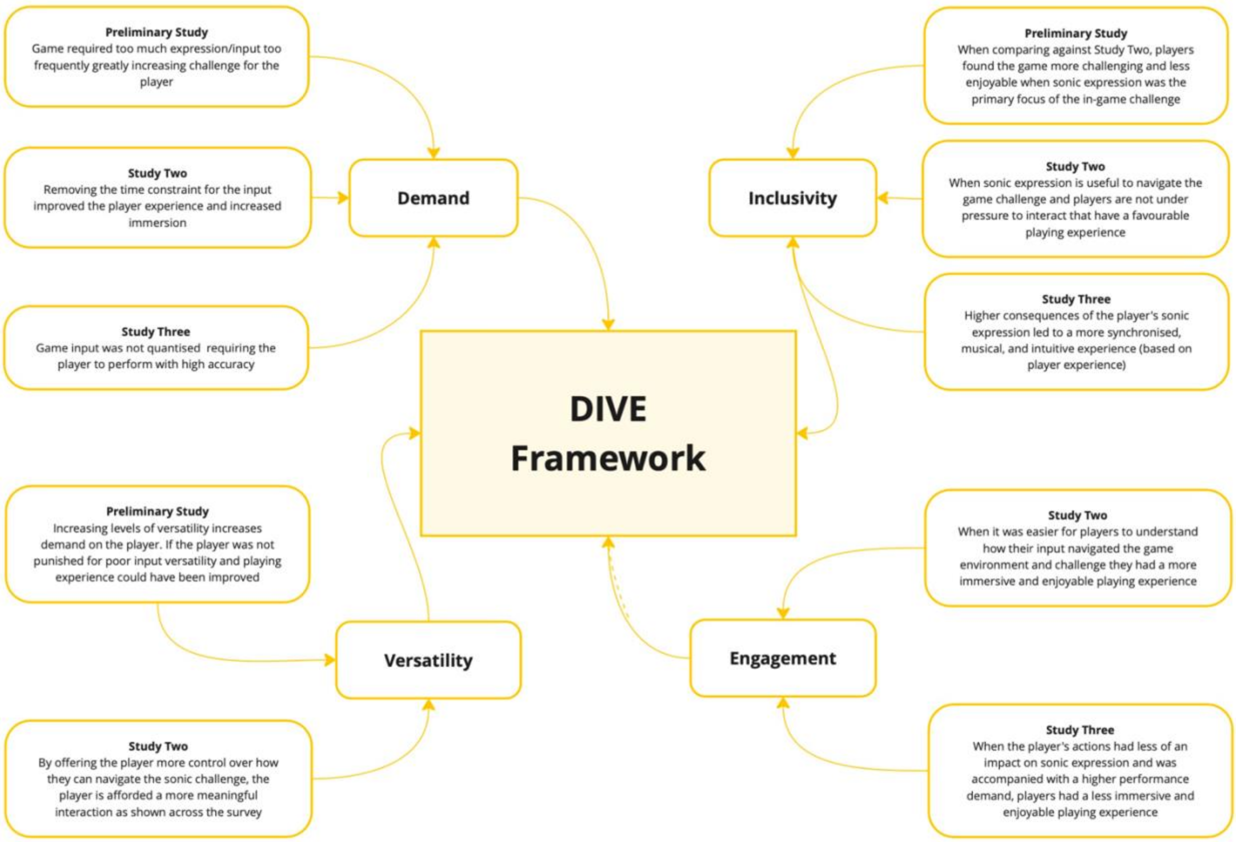


Figure 48: Diagram of DIVE Framework

Demand

The findings from all three studies show a clear relationship between the required level of sonic expression needed to successfully perform the in-game challenge and the immersion of the player. In the first and third studies. the games required the player to frequently perform in a time-sensitive manner. In the first prototype, the performance demand was too high. Players were required to apply spatial reasoning in a musical arrangement, with only an abstract representation of the music space to guide them. In addition, they also had to frequently express themselves and trigger new sonic events to avoid losing points. The game A prototype for Study Two solved the performance demand of the former challenge by removing the time limit. This alleviated the burden of performance pressure for the players and allowed the players more time to make their own actions. Addressing this demand in a more meaningful interaction enabled a higher level of immersion. This is evident when comparing the in the IEQ 'Challenge' factor findings of the three studies; where the Preliminary Study VR ($M = 3.25$) and the Study Three ($M = 3.231$), compared to the findings in Study Two ($M = 3.333$) (as shown in Table 11 on page 91, Table 22 on page 122, and Table 34 on page 151). These results can also be supported by the difference in how challenging the players reported the games were in the first part of the survey for in the Preliminary Study game A ($M = 4.33$) and Study Two game A (3.77). Subsequently, we can instruct game designers to be wary of making sonic expressions time-sensitive, as it encourages the player to rush their expressions, as opposed to making a meaningful, well-informed expression. This is important for encouraging high levels of immersion. This example highlights how the factor of demand relates to inclusivity in this framework. Another issue in the performance demand found in the Preliminary Study was the value of expressive interaction, as the players were deliberately given no way of deleting input. This meant if they wrongly timed an input, they would have to develop their rhythm around the incorrect input, or else they would be punished for a lack of coherency. This was also addressed in game A of Study Two, where players were given the option to delete any placed note. This highlights how demand relates to the factor of versatility in this proposed framework. By considering Versatility in our game design, we can lower the demand for sonic expression, positively impacting the immersive experience.

In a similar fashion to the Preliminary Study, the demand for interaction in the third study was time sensitive. Unlike the prototype in the first study, where player expression was quantised to the nearest 16^{th} musical note relative to the backing track, the interaction in the third study game was not quantised. This meant that players had to perform with high levels of accuracy to synchronise their input in time with the backing track. Although this challenge had fewer components than the one found in the Preliminary Study (this challenge had two components: accurate timing and the use of motor skills/spatial reasoning to target and shoot the blocks. The Preliminary Study challenge also required the player to consider the coherency of musical arrangement); the demand of sonic expression on the players is still too high. This is especially prevalent when triggering non-diegetic sound, particularly for players who do not have any musical experience, as when they timed their expression incorrectly, it disrupted their sense of immersion. This is reflected in the IEQ factor of 'Cognitive Involvement' when comparing this factor between Study Two (M = 3.533) in Table 22 on page 122 and Study Three (M = 3.281) in Table 34 on page 151. Where possible, game designers should lessen the importance of micro-timing. While it is important for the player to perceive a corresponding sound for their action (Tanaka, 2010) (Poepel, 2005), designing challenges where this is a core component is not ideal for immersion. This demonstrates how demand is connected to the other domains of Inclusivity, Versatility, and Engagement. Lastly, it is important to have some identifiable level of demand. This is reflected in the findings between each game in Study Two, where increased versatility in game A increases demand and enhances immersion (game A M = 3.329, game B M = 2.968, as shown in Table 22 and Table 23 on page 122). In the first section of the survey (found on page 116) on average, participants reported a slightly more intuitive, responsive, and controllable experience in Game B, where demand was lower (M = -0.056, SE = 0.822, $t(29) = -0.370$, $p < 0.05$, effect size d = 0.822).

Along with the findings between each game in Study Three, a difference in relevant background experience impacts the perceived discernibility (Ermi & Mäyrä, 2005), which negatively impacts immersion. Identifying and subsequently alleviating excessive demand is vital for designing engaging and immersive challenges.

Inclusivity

The findings from the first and second studies show if we are aiming for our game to deliver an engaging and immersive experience, we should utilise sonic expression as a method for navigating the in-game challenge, as opposed to being the primary focus of the challenge. It is important that the use of sonic expression is not a crux to the in-game challenge, and rather, the challenge is inclusive of sonic expression. The choices made in sound design that form the sonic expression should frame the challenge. This is clear in the results of the third study, where there is evidence that the approach taken in sound design influenced how the players engaged with the in-game challenge (IEQ 'Challenge' factor – game A M = 3.462, game B M = 3.231 as shown in Table 33 and Table 34 on page 151). This is also supported by the findings of the first section of the survey on page 144, where on average, participants players reported that in version A, the audio-visual experience felt more synchronised, that it felt more like a musical experience, was intuitive, was responsive, that they focused on syncing their input to the backing music, and that by making mistakes it broke their focus more (M = 0.187, SE = 0.677, t(25) = 1.407, $p < 0.05$, with an effect size d = 0.677). In relation to this, the in-game challenge itself must be clear to understand. It should be easily understood by the player and framed contextually in a manner that is not reliant on their perception and cognition of sound or music. As seen in the puzzle prototype in the Preliminary Study, when the in-game 'problem' to solve is intrinsically tied to sonic expression, it increases the difficulty of the challenge and negatively impacts the playing experience. This is also evident when comparing between the two different game versions of Study Two where all immersion factors increased between game A (M = 3.329) and game B (M = 2.968).When the problem can be solved through expression while maintaining a separately identifiable game challenge (such as navigating from a start point to an endpoint, as seen in the prototype for the second study), then the use of sonic expression will positively impact the playing experience. In both the first and second (Game A) study prototypes, the in-game challenge was designed to be open; players could solve and win the game based on their own decisions. There was not a predetermined order of inputs to match. Consequently, the difference in playing experience is due to the puzzle game challenge having a higher demand for comprehension of music, asking players to make internally (to the player's previous input, not predetermined by the game designer) coherent decisions than that of the 3D platform game challenge.

It is important to note that when designing gaming challenges inclusive of sonic expression, game designers should avoid using sonic expression passively or ancillary to the in-game challenge. As discussed in the Demand domain above, the sonic expression must remain meaningful (Ermi & Mäyrä, 2005) to aid immersion. The findings from the third study reflect this in relation to Inclusivity. In the prototype for Study Three, the sonic interaction was found to lack meaning by the players depending on their prior musical (and music software) experience (as reported in the Study Three evaluation on page 153, where those with musical experience reported that it felt less like a musical experience (M = 2.533, SD = 1.060) than those without musical experience (M = 3.182, SD = 1.052)). This was due to the symbiotic sonic challenge (the embodied interaction, which relates to sonic expression and in-game challenge design) lacking sufficient degrees of Inclusivity. While music games will never meet the same capacity for sonic expression as music production software, game designers can still make engaging and immersive experiences by meaningful application of sonic expression to navigate the game challenge. The prototype for the first study suffered equally but for a different reason. Rather than a lack of meaningful sonic interaction, players found the challenge in the first study prototype to lack meaningful context. Players had an abundance of sonic expressions at their disposal but found the challenge difficult to contextualise. By designing clear challenges inclusive of sonic expression and contextualising the sonic expression within the challenge, game designers will be successful, as reflected in the findings of the second study. This is heavily connected to the domain of Versatility.

Versatility

By offering the player more control over how they can navigate the sonic challenge, the player is afforded a more meaningful interaction. This provides that there are sufficient degrees of inclusivity and that the Engagement of sonic expression is continuous (relative to the game challenge). As seen in the findings for Study Two (where all immersion factors increased between game A (M = 3.329) and game B (M = 2.968) and where players reported higher levels of motivation for game A (M = 3.5) compared to game B (M = 2.767), along with higher of enjoyment in game A (M = 3.367) than in game B (M = 2.767), and were more likely to replay game A (M = 3.4) as opposed to game B (M = 2.5) as reported in the Study Two evaluation on page 121), this is significantly beneficial for creating highly motivating and immersive experiences. Similarly, the findings from Study Three also reflect this. When sonic expression was limited, players did not feel motivated, and immersion was negatively impacted. This was apparent in the results of the IEQ for players that had previous experience of music production software, where they reported a lower sense of immersion (M = 3.203, SD = 0.343) than those without (M = 3.611, SD = 0.472). In both instances, Demand is also affected, as increasing levels of Versatility impact the Demand on the player. This is also evident in the results of the Preliminary Study. These results serve as a warning that a versatile game challenge must still have the challenge component communicated effectively. However, it is worth noting this particular challenge lacked the ability to delete input, which would have improved the game's Versatility. Likewise, if the players were not punished for lack of input, this would have also improved the Versatility of this game. This emphasises the relationship between Versatility and Inclusivity. Subsequently, game designers must be aware that any change they make to improve levels of Versatility, such as increasing the range of expression (e.g., dynamics of a triggered impulse), will have an impact on the Demand domain and will require suitable context for Inclusivity.

This makes Versatility difficult for game designers to address. Designing engaging game components that have a multitude of potential outcomes is more time-consuming than designing a component that has a singular prescribed outcome. The Game A prototype for the second study made use of procedural content generation to facilitate Versatility in the game. By using player sonic expression to modify generated content, a novel series of platforms (and enemies to avoid) was generated for the player to navigate. While this approach was successful in this instance, game designers must be aware that procedural content generation is not a panacea for Versatility in games, as procedural content generation often leads to less immersive level design compared to that of a human-designed level (Connor et al., 2017). The generated content must still be meaningful for Inclusivity, and the game designer must be careful not to negatively impact the challenge by using it. Additionally, designers must be wary of any unintended consequences or behaviours derived from the algorithms used for procedural content generation. This requires extensive testing, adding an additional burden on development time.

Engagement

Engagement can be considered the dialogue between challenge and sonic expression. This domain pertains to how well the designed expression communicates within the game. It also relates to how the corresponding interaction engages with the in-game challenge. The use of sonic expression and the corresponding interaction used to modify the sound design of the game works best when there is a clear impact but does not disrupt the challenge. Game designers need to clearly communicate the context of the sonic expression relative to the game challenge. High levels of Engagement result in a clear understanding of how the player's expression relates to the challenge. When sonic expression is continuous in relation to the challenge, the player can quickly form an understanding of how their expression can be used to navigate the game challenge. Consistency in interaction and behaviour of the game is imperative for the playing experience, as consistent behaviour helps players predict outcomes, which motivates further experimentation (Sale & Zimmerman, 2003). This is likely reflected in the cognitive involvement of players and positively impacts the overall playing experience and sense of immersion. This can be best shown by comparing the results of studies two and three.

In both games, the player's expression changed non-diegetic sound-design in the game. However, the reported playing experience and sense of immersion were lower in Study Three compared to that of Study Two. This was in part due to lower levels of Engagement. In Study Three, if players did not time their input correctly in Game A, they would trigger the bass notes out of sync with the overall backing track. Despite disrupting the intended non-diegetic sound in the game, this had no impact on the behaviour of the game challenge. Only if the block was missed would it have a resultant impact on the in-game challenge. This can be considered low Engagement (while having a high Demand, as discussed above). This is because there was no meaningful relationship between the sonic expression and the challenge. This is apparent when comparing the IEQ factors of 'Control' and 'Challenge' of game B of Study Three (Control M = 3.02, Challenge M = 3.131) against game A of Study Two (Control M = 3.386, Challenge M = 3.333), and the overall difference in the player experience survey results for these two games (Study Three M = 3.239, Study Two M = 3.302). If the player shot the target at the earliest opportunity or the latest opportunity before it damaged them, the result for the in-game challenge would be the same, regardless of whether the expression (rhythmic timing) of the shot was different. This is vastly different to the Engagement in Game A of the second study. In this game, when players placed their chosen token to trigger a note, they were presented with a clear impact that was consistent for every interaction of that type of note. Importantly, it also included noticeable consequences for the completion of the gaming challenge. Additionally, when players make progress through the game (after collecting tokens), the backing music changes. This change is not instantaneous;

it waits until the current dynamic loop of music (Collins, 2007) has finished. By transitioning the backing music relative to the properties of the music, as opposed to instantaneously, Engagement is less disruptive and remains meaningful. Many non-linear sound-design approaches (Collins, 2007) that are already employed by game designs are appropriate for this task. By making slight modifications to either audio source, DSP behaviour, or triggering, these approaches would be suitable for yielding high levels of Engagement and practical for sonic expression in game design.

Summary of DIVE Framework

While this framework is experimental, by addressing each of the four domains (Demand, Inclusivity, Versatility, and Engagement) within game design, this researcher anticipates the resultant game would yield higher levels of overall immersion. This would likely positively impact the overall playing experience. Each individual domain identifies distinct factors relevant to immersion and connects existing research on interactive sound, immersion, and game design. The findings of this research provide evidence for this framework and highlight how each domain is interconnected to one another while still maintaining a clear distinction. A summary of games (including the research prototypes) and how they are analysed using the framework can be found in the table below.

Table 35: Games Analysed with DIVE Framework

Game	Demand	Inclusivity	Versatility	Engagement
Splinter Cell: Chaos Theory (Ubisoft, 2002)	While generating sound (via your character tapping a wall) can be used to alert nearby NPC, it is not needed to complete the challenge and has a low/binary use of sonic expression. **Low Demand.**	Sonic expression is not a crux but is entirely ancillary to the in-game challenge. The player has very limited expression at their disposal. **Low to Moderate Inclusivity.**	The use of sound is only a limited method to navigate the challenge, which is not explicitly a sound challenge. **Low Versatility.**	While the player is very limited in how they can express themselves in sound, what actions are available are very clear and have a predictable response. It can also be used in a variety of different in-game scenarios to effectively navigate the in-game challenge. **Moderate Engagement**.
Guitar-Hero (Harmonix, 2005),	The demand is limited within one dimension – rhythmic sequence matching. However, this ranges from low-effort to high-effort	Sonic expression is a crux, and the challenge can only be completed by explicitly following the challenge. **Low Inclusivity.**	The method to navigate the in-game challenge is singular and not open to be solved in different ways. **Low Versatility.**	While the player is limited in how they can express themselves in sound, the action available to them is clear and directly responsible for

	sequences. **Low to High Demand.**			completing in the in-game challenge. **Moderate to High Engagement.**
Rez: Infinite (Monstars, 2015)	The demand on the player is relatively low as there are little to no consequences for the player missing their input, providing they are still within the window before the enemy hits them. The audio-visual has a comfortable window for synchronisation. **Low Demand.**	Sonic expression accompanies the input necessary to complete in-game challenge but not explicitly necessary (although does signpost input well). **High Inclusivity.**	The use of sound to navigate the in-game challenge is closed and although there are slight variations of sound that can be triggered, the player has limited input. **Low to Moderate Versatility**.	The actions available to the player are limited but clear and do not disrupt the challenge. **Moderate to High Engagement.**
Preliminary Study - Boxel (A and B)	Overly performative in the sound domain. The players could make use of the limited visual elements for appropriate input, but the games largely depended on their musical intuition or understanding of music theory to be successful. The challenge was time sensitive too, which also raises the demand on the player. **High Demand.**	The use of sonic expression was a crux. Although the game challenge could be solved in a number of a different ways it still required musical understanding. While the challenge could be solved without this, it was significantly harder to do so**. Low Inclusivity.**	The method to navigate through the sound-challenge space was open. The challenge could be solved in a variety of different ways, it was dependent on the players previous input not prescriptive. **High Versatility.**	Individual actions for generating sound were continuous and clear. Although players could need remove unwanted input. Additionally, as coherency is not checked against a fixed value this could be unclear to players (especially those who are inexperienced). Low to **Moderate Engagement.**
Study Two – Red is Dead (A)	Sonic expression used to complete in-game challenge but not overly performative as there was no time pressure to place notes (or specific position to place them). **Moderate Demand.**	The challenge is not dependent on the use of sonic expression, but the player has a range of expressivity. **Moderate to High Inclusivity.**	The choice of how the player can navigate the game environment and challenge is open**. High Versatility.**	Sonic expression is continuous, clear, and players can remove their previous input if necessary. **High Engagement.**
Study Two – Red is Dead (B)	There is no choice of expression as it is predetermined (although there is no time limit) **Low Demand.**	The challenge is not dependent on the use of sonic expression and the player has no agency regarding what sounds are triggered. **Low Inclusivity.**	The player has no choice how to navigate the challenge (although they are free to navigate the in-game environment as they wish). **Low Versatility.**	The use of sound is clear and continuous, but actions are very limited. **Low to Moderate Engagement.**
Study Three – Grail (A)	The game was overly performative, and players only had a small amount of time successfully complete their input before taking damage. **High Demand.**	Players were often negatively impacted when they mis-timed their input out of sync with the backing music. Expression was also limited to the predetermined arranged notes. **Low to Moderate Inclusivity.**	The players were limited in how they could interact sonically with the game challenge. They were only responsible for the micro-timing of notes. **Low to Moderate Versatility.**	While player experience was clear and continuous, incorrect input greatly impacted Engagement, as it interfered with navigating the challenge. **Low Engagement.**
Study Three – Grail (B)	The game was overly performative, and players only had a small amount of time successfully complete their input before taking damage. **High Demand.**	The audio-visual experience was less impacted by syncing issues compared to the A version above. However, overall expression was limited further as players only generated a sound effect. **Low to Moderate Inclusivity.**	Players were greatly limited in how they could express themselves in game. They could only trigger a sound effect. **Low Versatility.**	Players were less impacted by their sonic expression as their (limited) expression did not interfere with the audio-visual experience. **Moderate Engagement**.

This framework is needed as existing frameworks concerning sound design in games follow traditional approaches originally found in film studies (Huiberts & van Tol, 2008), which do not address the key relationship between sound and in-game challenge design for games that utilise sonic expression. Further research should be conducted to validate this conceptual framework in the future; further practice can use the framework for its design.

Contribution

As this research is practice-based, the contribution of this research is two-fold. The first contribution is the novel sound design approaches and innovative game mechanics as developed and tested for each of the three studies. Although each prototype was designed around a typical gaming paradigm (puzzle, platformer, rhythm game), each game contains at least a single, unique game mechanic that is not found in any other title or research project. Five of the six prototypes were VR games. The first prototype for the Preliminary Study features a music game with a novel in-game challenge, which involves developing a coherent musical rhythm. Unlike existing music games that are reliant on strict rules giving a limited scope of expression (Pichlmair & Kayali, 2007), the game does not feature any strict sequence matching to award points. Instead, the puzzle is solved, and points are awarded for how well players 'match' input to their previous interactions in the context of coherent rhythmic development. The game challenge requires deviation (the player cannot simply input the same ostinato repeatedly) and uses their expression as a defining feature of the game challenge. It is dynamic and has no strict sequence to follow, and thus, each playthrough will yield a different sequence to be matched against (providing the player expresses themselves differently each time). This is a highly innovative gaming mechanic, as the use of sonic expression scarcely exists in any capacity in sound design or in game design, let alone as a core gaming challenge.

The prototype for the second study also featured novel gaming mechanics and game design features. In the Game A prototype, to successfully navigate the game task, players were required to use sonic expressions. By taking a similar technical approach found in the first study prototype (classifying player expression based on note distribution and self-similarity across metric layers), this researcher developed a 3D platform game which required the player to use sonic expression (in the form of developing a sequenced drum beat) to generate the platforms required to reach an end goal. To make an expression, players had to collect tokens. When players used their collected tokens, the game would procedurally generate a new row of platforms that the player could utilise to get closer to the end goal. Each set of platforms generated had slightly different properties, dependent on what type of sample was triggered, with its form dependent on its relative position to the other notes sequenced by the player. Depending on how coherent the player was developing their rhythm, the procedural content generator would place the next collectable closer or farther away from the player. Additionally, each set of newly generated platforms would come with a moving playhead enemy to avoid; the form and movement speed of these were also determined by the coherency. Another novel feature of this prototype was the sound design approach. The sound design for the game was generated 'live' at run-time. The game featured multiple parameterised synthesisers, with each synthesiser having certain parameters controlled by player movement and interaction. One synth was used to trigger note intervals to be triggered upon player collision with the enemies. This approach is highly experimental and, to the knowledge of this researcher, not found in any commercially released game.

The third study featured a prototype (Game A) rhythm game where the player interaction triggered part of the non-diegetic soundtrack. Players were required to perform with highly accurate rhythmic timing to experience the intended soundtrack. This approach is unique, as it deviates heavily from how a traditional rhythm game presents the non-diegetic sound in its sound-design approach. By increasing the importance of the player's accuracy in timing, the game afforded more creative agency to the player. As covered in the literature review on page 37-38, existing rhythm games offer little in the way of creative agency, frequently have very low levels of player expression, and have a static gaming challenge, i.e., the required input sequence is the same every time the level is played. While the challenge in the prototype is the same in each playthrough and offers a similarly static gaming challenge, it does offer some creative agency. While freeform games exist that do allow for increased levels of player agency, these all lack an explicit in-game challenge. This game had both an in-game challenge and some degree of creative agency. Additionally, relinquishing control over the overall quality of arrangement (synchronised sounds) of the soundtrack is a highly novel approach to sound design.

The second contribution is the development of the conceptual DIVE (Demand, Inclusivity, Versatility, Engagement) framework for informing game design. By making and using several playing experience surveys, in conjunction with the use of the validated IEQ for immersion, this research has produced novel findings in the relationship between sonic expression, in-game challenge, and immersion. Future researchers will now have a firmer basis of understanding the role sonic expression can have in immersion and the overall playing experience. Additionally, these findings have been used to develop a conceptual framework for the use of sonic expression in game design. This has strong practical implications for the future development of music games in sound design and for the broader field of game design. By addressing each domain of DIVE during development, game designers will positively impact the player's sense of immersion and the overall playing experience. Future research in this field using DIVE will challenge the initial assumptions found in the framework. This will lead to an increased understanding of the role sonic expression has in immersion and will inform the development of the framework itself.

Reflection on Methodology

As practice-based research, the methodology for this research bridged the gap between game design research and game design practice. By limiting the variables between each version per study, the research successfully measured the impact value of that variable. This was helped by limiting the scope and overall design of each prototype. In practice, games are multi-faceted and often feature a high number of designed elements. Many games include a variety of game environments for the player to traverse, along with interactive narrative and a highly detailed audio-visual experience. By reducing the game design down to only the necessary core components required for the game paradigm and sonic expression, this approach mitigated the potential difficulty of identifying the influence factor for our measured playing experience. This benefitted research.

Throughout each study, attempts were made to maintain a consistent set of questions within the post-survey section pertaining to the overall playing experience. This was difficult to accomplish due to each study making use of a different gaming paradigm (puzzle, platformer, rhythm game). Each of these paradigms has vastly different characteristics and features different components, making survey design hard as it was difficult to meaningfully address specific shared components. However, through multiple iterations over the course of this research, the overall survey was improved significantly. The removal of some questions and overall shortening of the survey experience helped in this regard. The use of the validated IEQ to measure immersion was beneficial, as the five immersion factors provided excellent resolution and were straightforward to interpret. Another limitation of this approach was that while choosing a different paradigm for each study effectively broadened our understanding of game design, it was not conducive to providing nuance. By iterating and testing a single paradigm, this would yield a more detailed understanding at a granular level. It would be wise for future research to follow this method.

Despite the uncertainty and significant challenges that the COVID-19 pandemic imposed on the initial study design of the research, the overall study design was improved by the necessitated changes. By catering for remote participation, it crystalised game instructions delivered to players, imposed further technical limitations on the game (which benefitted design creativity and forced a greater limited scope in game design), and increased the pool of potential participants as they were not needed to attend a specific physical location. While this came with a necessary extra demand for player verification, this was not difficult to overcome. The change in the development environment from Unity (C#) to JavaScript/HTML with A-Frame and MaxiInstruments was much harder to overcome. This change, in combination with the change of platform delivery from an Oculus Rift DK to an Oculus Quest 2, was the largest limiting factor in this approach. This will be discussed in more detail below. However, the change to deliver the game and survey online to participants was a net positive for the overall study design within this research.

Limitations

As discussed above, the use of multiple gaming paradigms potentially limited the level of detail found across the research. While the second and third studies both had a healthy number of participants, it is prudent to acknowledge that the Preliminary Study had significantly fewer. This limits what can be interpreted out of these findings independently from the other two studies. However, this is not a major concern as appropriate data analysis methods were used for the small sample size, and the results were found to be significant and not due to chance. Although the conclusions that can be drawn from the Preliminary Study are limited, the knowledge gained from the study was vital in informing the design of the following studies. Subsequently, the Preliminary Study contributed meaningfully to the overall methodology of this research. While the second and third studies had enough participants to provide conclusive results, increasing the number of participants for these studies may yield further insight as well. Especially if the participants represented a broader demographic spread; however, this would be difficult to accomplish without a significant promotion and social media reach. Another limitation related to the overall study design was the limited playing time of participants. The relative experience shaped a noticeable impact on the player's immersion, and it would be worthwhile to observe this change by periodically surveying them over larger intervals of time. Observing the player's experience using qualitative research methods would have also supported the quantitative data and potentially codified the results in greater detail.

As discussed in the literature review on page 50, immersion can be measured in greater detail by combining the use of a questionnaire and objective data. The use of 'automatic event logging' is a long-standing practice (Hilbert & Redmiles, 2000) in HCI and provides key observations relating to the experience of the user (Nacke et al., 2008). Taking these methodologies adapted specifically for games Pagulayan et al., 2012) (Pagulayan & Steury, 2004) would have yielded more detail regarding the playing experience in relation to immersion. Unfortunately, this was not possible due to the time constraints of the development window, in conjunction with the limited technological affordance available in the second and third studies. Participant performance during the Preliminary Study was captured and informally observed but did not feature in the formal analysis of the playing experience. Addressing this limitation in the future would likely be beneficial.

The change of platform from an Oculus Rift Development Kit to an Oculus Quest 2 imposed several limitations. Chiefly, it limited the computational resources available for running the prototypes. To increase the accessibility of the games (and improve remote participant recruitment), the game delivery changed from application to within browser. This resulted in having to optimise each prototype significantly so it could run without significant issues affecting gameplay. This was time-consuming and resulted in minimising some of the features originally in the game, albeit not at the cost of the primary feature incorporating sonic expression. However, it is important to consider that these comprises may have resulted in the player missing potentially engaging features, ultimately limiting their experience. Another limitation caused by this change in technology was the lack of agency in developmental oversight. Frequent updates to the Quest 2 browser would result in stability issues for software written by this researcher and the external libraries written by others (A-Frame, MaxiInstruments). While the game software could be rewritten, changes to external libraries were out of my control (beyond identifying the issues and notifying contributors). Additionally, unlike C#, where developers can control memory allocation and garbage collection, this was not possible in JavaScript in the Oculus browser and had the potential to cause stability issues in performance. Likewise, a by-product of running the study remotely was that it was difficult to troubleshoot any performance issues; as such, it is hard to ascertain whether negative playing experiences were caused by playing experience or a non-related platform issue.

As this research is practice-based, identifying the limitations of game design is also relevant. The prototype in the Preliminary Study was primarily limited by the player's comprehension of the gaming challenge. If the relationship between sonic expression and points scored was communicated clearly, using additional graphics, this may have alleviated some of the issues experienced by the player. A secondary limitation of the prototype in this study was found in the screen-based version of the game. Players reported the gestural method of interaction was limiting their ability to perform. This issue would be hard to address without significantly altering the interface or game environment of this prototype. The most appropriate solution would be to divide the touchscreen into quadrants and use a tap/click as an input trigger. However, this would differentiate the game significantly from the corresponding VR version. The prototype in the second study was the game most limited by the change in technology. The game had a high computational overhead and required refactoring and optimisation. This resulted in stability issues and the occasional glitch. Although removing some of the sophisticated logic for item placement helped, the deliverable playing experience was impacted by the occasional glitch. In attempts to mitigate this, players were warned the game was experimental. The prototype for the third study was limited similarly, although the use of chunking and preloading data mitigated most of the game's instability. Subsequently, computational performance was mostly acceptable, with the occasional drop of frame when there were many targets on screen (towards the end of the level). Although the sound design approach was deliberate in this study, results could potentially be improved by addressing the quality of synchronised input. Quantising input may have yielded better results as 'a disjointed score generally leads to a disjointed playing experience, and the game may lose some of its immersive quality' (Collins, 2007).

Practical Implications and Future Research

As covered earlier in this chapter, this research contributes to both the practice of game design and future academic and industry research. The conceptual DIVE framework can serve as a guide for other game designers to follow, aiding designers to make highly immersive, engaging experiences that utilise sonic expression. This is especially important for those new or inexperienced in sound design as there is little existing theoretical support for them (Alves & Roque, 2011) (Collins, 2015). The prototypes developed for the studies can also serve as tangible inspiration for novel game design. The novel gaming mechanics found within each prototype can serve as a basis for future experimentation and can be adapted for different gaming scenarios. Other game designers have an example to follow, hack, and reappropriate as they see fit. Furthermore, the three prototypes will be developed beyond their current state within my own personal practice as a game designer. The findings from Study Two show that the Game A prototype was highly successful at providing an immersive and engaging playing experience. This gaming mechanic will be developed further following the guidelines laid out in the DIVE framework, with the aim of incorporating it into a rich game environment.

The findings of this research have revealed unanswered questions that could serve as a basis for future research. Studying the impact of how prior experience influences immersion would provide valuable insight into potential criteria for cognitive involvement and challenge immersion. It would also elucidate the role of novelty in player engagement. The findings from Study Three suggest there is a sufficient requirement of challenge that needs to be met, which can be influenced by transferrable prior experience before players feel immersed. To what extent this occurs would be useful for game design research. This is especially important for challenges pertaining to sonic expression, as identifying other transferrable experiences may lessen the demand for interaction. Further study on the role of agency in our sound interactions would be beneficial to challenge the findings of this research. Studying the impact of interaction methods that afford a wider range of sonic expressions, such as the use of dynamics, would yield relevant research that either strengthens or weakens the findings here. Whether there is a measurable difference in immersion between different sound elements, such as pitch or timbre, would also be an interesting parameter to measure. Determining if there are suitable thresholds for the agency when manipulating these characteristics would prove useful for future designers and would help further our understanding regarding the role agency can have in immersion. Another area of future research should be exploring the disruptive capabilities of sound design. As found in the third study, certain sound design approaches can have a negative impact on immersion. It would be beneficial to study these in more detail, determining where the lower limits of engagement occur. This could be explored by breaking the continuity of expression through a variety of different methods and measuring the impact on immersion.

These could be distilled into three new future research questions:

- To what extent does prior experience impact cognitive and challenging involvement in immersion?
- What is the relationship between dynamic range, sound elements, and immersion?
- The role of continuity in immersion.

Lastly, all three research questions would be suitable to explore using the conceptual DIVE framework herein. This would further develop, expand, and validate the framework. Addressing each of these future research questions will expand the contribution made within this research.

Conclusion

By interpreting the findings from the three studies undertaken, we have addressed the three key questions of this research. It was found that sonic expression impacted the playing environment and the player's sense of immersion in a variety of different ways. The findings established that increasing agency in sonic expression had a significant positive impact. If the necessity of discernible meaning were missing in the interaction, it would result in a negative impact. The results also show that sound design plays an important part in player enjoyment and engagement. This was to a noticeable extent, where the findings show a nuanced relationship between sonic expression and the five immersion factors (cognitive involvement, emotional involvement, real-world dissociation, challenge, and control) that were measured for each study. The research also found that increasing the depth of sonic expression did produce measurable changes in immersion. Harnessing sonic expression within the in-game challenge uncovered a complex relationship between the use of sonic expression and playability. All the findings were used as a basis for developing an experimental conceptual framework, specifically for the use of sonic expression in game design. The DIVE (Demand, Inclusivity, Versatility, Engagement) framework connected four domains that are utilised in game design, inclusive of sonic expression. By carefully considering each domain and following the guidelines provided, game designers will increase levels of immersion in their game and improve the overall playing experience. This discussion of results found several clear research contributions and areas of practical innovation, outlining the positive practical implications of this practice-based research approach. By reflecting on the methodology used, it was possible to identify the limitations of this research. Solutions for these limitations have been provided to aid future research, along with a set of future research directives, in the form of new questions to answer. These questions aim to expand the contribution found in this research, provide valuable literature for sound designers, and encourage further development in the use of sonic expression for game design.

Conclusion

Overview

As outlined in the introduction of this thesis on page 17, the use of sound to immerse the player in games is often a passive response to the player's in-game interaction. This practice-based research addressed this by utilising sound actively as a key variable within games, with the intention to positively impact player immersion and their overall playing experience. Through measuring the impact of sonic expression and identifying factors of immersion and the playing experience, a framework for the use of sonic expression in game design was developed. This was achieved by developing several unique prototypes, each featuring novel game design mechanics that were centred around the use of the player's sonic expression to navigate an in-game challenge. These were evaluated over a series of three studies, with each prototype concentrating on a different facet of game design while still encompassing the use of sonic expression and were expressed as a set of three research questions:

- To what extent does sonic expression impact the playing experience?
- Does increasing the depth of sonic expression produce a measurable change in player immersion?
- How does sonic expression impact game challenge, and does this affect immersion?

Answering all three of these research questions was important to identify and inform all domains of the DIVE framework. By answering 'To what extent does sonic expression impact the playing experience?' the research highlighted key areas of game design that needed further exploration, as it identified which components of game design would be most effective for experimentation. It identified the Demand domain as it was found that sonic expression significantly impacted the challenge component of all prototypes, along with affecting control and the cognitive involvement of the player. Inclusivity was informed by the impact of sonic expression on the challenge component, recognising that sonic expression should not be a challenge in and of itself. The impact of sonic expression on the player's motivation, enjoyment, and desire to replay the game then shaped the Versatility domain.

Engagement further highlights the potential negative impact that sonic expression can have on the player. By establishing that a causal relationship between the depth of sonic expression and change in immersion, the experimental results of the studies strengthened the framework. This was especially important for strengthening Demand, Versatility, and Engagement. The findings show that if the Demand of sonic expression is too high, then it negatively impacts the playing experience. Likewise, if the Demand is too low, the playing experience is negatively impacted for some players. Versatility was also informed through the findings for this question, as it was determined that increasing the depth of sonic expression, when harnessed within an open-ended challenge, led to a positive playing experience. It was also found that if the depth of sonic expression is not sufficiently communicated within the game, it would negatively impact the player's Engagement. Explicitly answering how sonic expression impacted the in-game challenge and corresponding sense of immersion addressed the core gaming mechanic of the challenge. This reinforced all domains of the framework.

The resultant DIVE framework is invaluable as existing frameworks concerning sound design in games do not address the use of expression with sound. By deviating from other sound design frameworks that follow approaches originally found in film studies (Huiberts & van Tol, 2008), the DIVE framework can address specific game design issues arising out of the novel use of interactive sound.

Key Findings

The first study, 'Comparing the playing experience of VR and Mobile platforms when using sonic expression as a gaming mechanic,' established that there was an identifiable advantage in developing sonically expressive games in VR. It was determined that the technological affordance of VR over screen-based platforms was beneficial to enhance the playing experience (Screen-based vs VR $M = -0.3917$, $SE = 0.370$, $t(11) = 3.672$, $p < 0.05$, with an effect size $d = -1.075$). This was supported by findings from scoring the 31 IEQ (VR $M = 3.658$, Screen $M = 3.306$). The results from the Preliminary Study showed players found the game less challenging when in VR ($M = 3.167$) compared to playing on-screen ($M = 4.333$) and reported they felt significantly more in control in the VR version ($M = 3.833$) than the screen-based version (M = 2.083). Players also reported the controls were more intuitive in the VR version (screen M=3.35, VR M=4 Q6 Part 1). The results also helped inform game design approaches regarding how to integrate sonic expression into game challenge design.

The second study, 'Studying the impact of player agency over rhythmic sequencing and platform placement on the playing experience,' explored the impact of the player's creative decisions to navigate a typical in-game challenge. It was found that player agency can be leveraged to make novel and engaging playing experiences. By offering players more agency in their musically expressive decisions, players reported higher levels of engagement and immersion. This was reflected in the findings of the second section of the post-survey (Game A section two M = 3.28, Game B section two M = 2.830) and in the scoring of the IEQ (Game A M = 3.15, Game B M = 2.85). This was prevalent in the factors of player motivation and enjoyment. There was a substantial impact on player motivation (Game A M = 3.5, Game B M = 2.767), with a similarly significant impact on enjoyment (Game A (M = 3.367, Game B M = 2.767). Players also expressed a much higher desire to replay the game with increased agency (Game A M = 3.4, Game B M = 2.5). The study also highlighted that increasing agency led to players finding the game more challenging (Game A M = 3.77, Game B M = 2.9). This was also within the immersion factor of the challenge of the IEQ (Game A M = 3.333, Game B M = 2.625). The findings of this study can be utilised to inform future game design, highlighting how simple procedural content generation can be harnessed when using sonic expression, leading to an engaging game challenge.

The third study, 'Music game design and the impact of non-diegetic synchronised cues on the playing experience,' explored the significance of sound design decisions in relation to the player's interaction that was required to play the game challenge. It determined that, despite reporting a lower sense of immersion, players had a more engaging experience when their interaction was responsible for the non-diegetic sound. The use of non-diegetic musical sound for player interaction negatively impacted the specific IEQ immersion factors of cognitive involvement (Game A M = 3.449, Game B M = 3.568) and control (Game A M = 3.077, Game B M = 3.131) but impacts the factors of challenge positively (Game A M = 3.462. Game B M = 3.231). It also was found that immersion was impacted by the player's prior experience of the use of electronic music production software. Players with music production software experience reported a lower sense of immersion (M = 3.203, SD = 0.343) than those without (M = 3.611, SD = 0.472). However, despite reporting a lower sense of immersion, players reported a more engaging experience when their interaction triggered non-diegetic musical sound (M = 0.187, SE = 0.677, t(25) = 1.407, $p < 0.05$, with an effect size d = 0.677). Players also reported that when their actions generated non-diegetic sound to form part of the backing music, they focused more on synchronising their playing inputs. This highlights a nuanced relationship between player experience, in-game challenge, and immersion.

Answering the Research Questions

The key findings from these three studies provide valuable insight into the extent to which sonic expression influences players' immersion and engagement. Consequently, we can answer the original three research questions and provide sufficient understanding to develop the DIVE framework to guide future game development and research.

The first research question, 'To what extent does sonic expression impact the playing experience?' is addressed through each of the three studies. Each study highlights a different aspect of the multifaceted relationship between the use of sonic expression and the playing experience. The Preliminary Study laid the foundation for understanding how sonic expression contributes to the playing experience. This study found that the integration of sonic expression into gameplay enhanced the playing experience positively. However, it did stress that using sonic expression as a gaming mechanic can increase the difficulty of in-game challenges. This raised an intriguing question regarding cognitive involvement and the potential conflicts between auditory perception, cognition, and spatial reasoning in the context of gaming challenges. This potential conflict may arise when players perceive discrepancies between their expected mapping of sonic interaction and the actual in-game experience. The use of specific sound design choices and the technical affordance of in-game sound interactions may also play a role in this phenomenon.

In the second study, the research delved deeper into the relationship between player agency, sonic expression, and the overall playing experience. These findings indicate that increasing player agency significantly enhances the playing experience and immersion. Two key factors that contribute to this effect were identified. The first was that a clearer and more intuitively designed game challenge led to significantly greater player motivation and enjoyment. In contrast, when the challenge involved complex musical comprehension, players struggled (notably those without a strong musical background). The second was increasing player agency, which led to a broader range of sonic expression opportunities, fostering overall player engagement and immersion. This aligns with the notion that immersion in games involves a willingness to concentrate, which is essential for player engagement (Cairns et al., 2014) (Jennett et al., 2008) (Grodal, 2003) (Klimmt, 2003).

The third study explored how the quality of sound triggered by player interactions influenced the playing experience and immersion. The research revealed that changing the type of sound triggered, whether diegetic or non-diegetic, has a discernible impact on the different facets of immersion and the overall playing experience. This highlights the significance of sound in shaping the player's engagement and enjoyment. The findings challenge conventional sound design frameworks, emphasising that context and player perception can influence the effectiveness of sound design. The study suggested that gameplay that triggers non-diegetic music events, despite being marginally less immersive, can provide a more engaging experience. The findings from all three studies show that leveraging sonic expression as a gaming mechanic can enhance player engagement if it is integrated thoughtfully into the game's design. It shows the need for a nuanced understanding of sound's role in gameplay, and the potential for innovative sound design approaches to enhance player engagement. Answering the second research question can further this understanding.

The findings of the three studies can also address the second research question: 'Does increasing the depth of sonic expression produce a measurable change in player immersion?'. Each study contributed to the understanding of how different aspects of sonic expression impacted the player's sense of immersion, highlighting the technological affordances offered within the game design (Pichlmair & Kayali, 2007) (Kurtz, 1998) and the game design choices that influence immersion. Although the first study did not explicitly explore the depth of sonic expression, it did offer significant insight. It demonstrated that VR has the potential to provide a more immersive gaming experience compared to screen-based platforms. This was not solely attributed to the richness of sonic expression, but rather, it emphasises the role of embodied interaction in achieving immersion. These results align with the idea that interaction must meet player expectations for immersion to occur (McMahan, 2003) (Tanaka, 2010) (Poepel, 2005), highlighting the importance of technological affordance (Gibson, 1979) (Gaver, 1991). VR, with its sensory dominance and larger-scale interactions, inherently supports a broader range of expression, thereby increasing the depth of sonic expression. This leads to measurable changes in immersion.

The second study conclusively demonstrated that increasing player agency and expanding the impact of a player's expression leads to significant measurable changes in player immersion. Players exhibited heightened immersion across multiple factors of the IEQ, reinforcing the notion that changes in technological affordance provided by the game design (Pichlmair & Kayali, 2007) (Kurtz, 1998) impacted immersion. This effect was influenced by improved device mapping (Fels et al., 2002) and the functionality and depth of sonic expression offered to the player. These findings highlight the importance of transparent device mapping (Fels et al., 2002) (Gibson, 1979) (Gaver, 1991) and the player's ability to navigate the space between interaction and sound comprehension, which are essential for achieving immersion. By intentionally designing the game challenge to be open and conducive to a greater depth of sonic expression, players had the freedom to express themselves musically and chart their own unique path within the game. This was achieved by implementing simple procedural content generation techniques that dynamically adapted the game world to the player's sonic expressions. Such an approach led to a novel playing experience with higher levels of engagement and immersion, as players were compelled to be more attentive and invested in completing the game's challenges.

The third study explored the impact of different sound-design approaches that indirectly broadened the depth of sonic expression within the gaming experience. By contrasting diegetic sound effects with non-diegetic musical responses, this study stressed the significance of sound in shaping player immersion. Importantly, it challenged conventional sound design frameworks, emphasising the need for context and player perception in sound design. The findings revealed that the players' prior musical experience could negatively impact their immersion, particularly for those with experience of music production software.

This discrepancy in expectation between experienced players and the game environment may stem from differences in the complexity of the challenge and the non-trivial impact of player actions on the environment. As the two different sound-design approaches resulted in differences in the arrangement of in-game music, this impacted the players' immersion. While one approach allowed for greater expressive rhythmic input, players had limited time to make their interactions, and it made it considerably harder to synchronise their inputs with the desired musical outcome. This highlighted the importance of technological affordance (Gibson, 1979) (Gaver, 1991) in facilitating player interaction with the sound and corresponding game responses. Collectively, the findings from these studies provide a nuanced understanding of the impact of increasing the depth of sonic expression on player immersion.

The third research question was, 'How does sonic expression impact game challenge, and does this affect immersion?'. All three studies found evidence that the use of sonic expression impacted the in-game challenge. The Preliminary Study showed the difficulty of designing games where sonic expression serves as the primary in-game challenge. Players exhibited a strong desire to overcome the in-game challenge in both the VR and screen-based versions, demonstrating their motivation and engagement. However, players found the screen-based version considerably more challenging than its VR counterpart. This increased difficulty was largely attributed to the abstract nature of the game challenge, heavily contextualised in music and rhythm. The results suggest that certain game genres, such as puzzle games, may not be the most suitable for using sonic expression, at least when the challenge is contextualised solely in music. Despite the high levels of perceived difficulty, players reported good levels of engagement and motivation, indicating that sonic expression may still contribute positively to the gaming experience, even when the challenge is demanding. The second study revealed that granting players greater control over their sonic expressions led to an increased perception of the game's challenge level. Players found the game more challenging when they had the freedom to shape their expression. Despite the heightened challenge, players also reported a more immersive experience. This result suggests that the relationship between sonic expression, game challenge, and immersion is multifaceted. The findings underline the fact that not all gaming paradigms respond to sound in the same way, emphasising the importance of considering game genre when utilising sonic expression.

The third study highlighted that sound design choices significantly impact the perception of challenge and, consequently, influence immersion. When players' actions generated non-diegetic sound to form part of the background music, they became more focused on synchronising their inputs. This synchronisation could be attributed to the player's motivation to bridge gaps and align their actions with the game's auditory cues. The study also revealed that the players' prior musical experience could impact their immersion negatively, particularly for those with experience using music production software. This discrepancy in immersion likely stemmed from differences in player expectations and the perceived impact of their actions on the game environment. However, despite participants with musical experience reporting a lower sense of immersion, players overall found the version triggering non-diegetic sound to be more musical and a more engaging experience. The findings reveal a dynamic relationship between sonic expression, game challenge, and immersion. This shows that the impact of sonic expression on immersion is not a one-dimensional relationship but rather a complex interplay influenced by factors such as game genre, player agency, and prior experience.

Strengths of Methodology

The largest benefit of the modular evidence-based approach undertaken throughout this research was that it led to clearly identifiable outcomes in each area. It was straightforward to identify how sonic expression influenced the individual factors of immersion and the overall playing experience. The study design was deliberately simple to reduce potential confounders. By reducing the different gaming paradigms to a low-level abstraction and minimising the non-essential (in the context of testing sonic expression) components of the game, including the game environment, graphics, and narrative, it was clear that sound and the use of expression within the prototype was the primary component. By using within-group A/B testing, factors indicating change in immersion and playing experience factors could be clearly differentiated as the use of sonic expression between those two versions. This approach produced substantial evidence for analysis and evaluation and resulted in the findings being relatively clear to interpret. This meant the findings could answer the specific research questions successfully and directly addressed the overall objective of this research. Another strength of this study design was how rapidly it could be deployed to a potentially diverse pool of participants. After the change in platform from Oculus Rift DK to Quest 2, once the first game and study website had been developed, it was straightforward to use this as a template for the following study. The simple website contained everything needed for the participants and was clear and straightforward to understand, and it minimised the physical demand of finding appropriate participants. In conjunction with the continued use of the validated immersion questionnaire (IEQ), this approach can be easily replicated in the future and would produce relevant, comparable results that could help further define and validate the conceptual DIVE framework developed within this thesis.

Through studying the use of sonic expression within three different gaming paradigms of a puzzle game, a 3D platformer, and a rhythm game, this research provides a broad understanding of how sonic expression can be utilised. This strengthens the DIVE framework, as it provides a broader contextual foundation. Further research will help expand this understanding and identify areas for refinement. The straightforward guidelines found within the DIVE framework also speak to the strength of the approach taken. This is directly attributable to the prototype design methodology, the quality of findings as described above, and the success in answering the research questions. The three prototypes developed also clearly emphasise the practical value of this research. The novel gaming mechanics developed here each serve as tangible, practical examples of experimental game design research. In combination with the clear and concise findings, this highlights the strength of how taking a practice-based approach can bridge the gap between game design research and game design practice.

Lastly, the approach emphasises the untapped potential which current game design practice has for novel and innovative techniques. This research demonstrates what can be achieved through combining explorative experimentation and an evidence-based methodology. This is especially important, given game design research is still developing.

Limitations of Methodology

While the strengths of the methodology are evident in the overall success of the research, it is important to identify and discuss the limitations to further reinforce a future approach. As outlined in the introduction, the scope of the game genre was limited to the specific predetermined gaming paradigms and would not necessarily generalise to all game genres. This is potentially the most prominent limitation of the approach undertaken. The use of these specific gaming paradigms limits the level of detail and nuance in the research. While it broadened the understanding of the impact of sonic expression, focusing on a single paradigm could have provided a more detailed examination in future research. Addressing this would be a significant change to the research structure, and ultimately would yield the equivalent limitation in reverse; if the study was designed to test a single gaming paradigm, the most prominent limitation would be the lack of breadth.

When conducting this research, it became apparent that the results could be more conclusive through several changes made to the study design. The third study's results highlighted the impact of previous relevant experience of the participants. In response to this, the findings would likely be more conclusive if players were observed over a longer period. Future research could benefit from periodic surveys over extended durations. In addition, the use of qualitative research methods would likely complement the quantitative data. Similarly, combining the use of the immersion question (IEQ) and objective data collection methods such as 'automatic event logging' (Hilbert & Redmiles, 2000) would provide a higher resolution of data for studying immersion. Unfortunately, this would likely have an impact on the development of the prototype and, depending on the platform, a potential impact on game performance.

The shift from using an Oculus Rift DK to Oculus Quest 2 imposed a serious impact on computational resources. Developing specifically for the Oculus Quest 2 web browser imposed a significant limitation on the game design. While A-frame performs well in-browser, sound does not. The inability to manage resources directly was an issue and may prove detrimental for other gaming paradigms or for testing games with a higher number of resources (or the complexity of the game). The only viable solution was to extensively optimise the game, which came at the cost of removing features. In a related issue, frequent updates made to the browser by developers outside of this researcher's control also caused issues with development. Future studies using the same methodology may also be impacted in the same manner. Lastly, the remote study design had many benefits but also had some noticeable limitations. Identifying any technical issue was incredibly difficult unless the participant proactively voiced an issue. This was then very difficult to troubleshoot remotely, with only generic solutions available (such as refreshing the page, closing, re-opening the browser, turning the device off and on, or doing a factory reset). Without the ability to accurately monitor the participation, it was difficult to confirm that participants had played without any issues.

While all these limitations have possible solutions to mitigate the issues, many are largely unfeasible within the constraints of academic research. The study design could take place over a longer period, but that would leave many months without results. A different delivery method could be used, but then it would impact the level of accessibility of the study and, depending on the platform, could be very costly.

Implications

This research carries significant implications for both game design practice and future research. The development of the conceptual DIVE framework offers game designers guidance regarding the use of sonic expression. This framework serves as a practical tool for game designers, especially those who may be new or inexperienced in sound design. It provides much-needed theoretical support in a domain where resources are currently limited. Game design practice is also supported by the development of the prototypes themselves. These prototypes can serve as a source of tangible inspiration for innovative game design, and the novel gaming mechanics discovered within these prototypes can be adapted and reappropriated for a variety of scenarios. The Game A prototype for the second study showed noteworthy promise as a highly engaging game concept and will be developed further by this researcher. This will be guided by the DIVE framework with the aim of integrating it into a richer game environment. In turn, this will also help to strengthen the framework. Likewise, the ongoing development within this researcher's own practice as a game designer underscores the practical impact of this research and reaffirms that this practice-based research has provided examples to follow, modify, and build upon to create engaging playing experiences.

The findings also have significant implications for future research. By addressing the main objective and determining the impact that sonic expression has on the player's experience and the player's sense of immersion, we have identified further questions that could expand our understanding. The findings uncovered unanswered questions concerning the influence of prior transferable experience on immersion. This poses interesting avenues for future research, with the opportunity to further strengthen game research methodology. Investigating how prior experience affects cognitive involvement and immersion, along with the role of novelty in player engagement, could highlight specific criteria for cognitive involvement and challenge immersion in gaming. Likewise, further research on the role of agency in sound interaction is warranted. Exploring the impact of interaction methods that afford a broader range of sonic expressions, such as dynamics, could deepen our understanding of how agency contributes to immersion. Determining measurable differences in immersion based on sound elements inclusive of a range (such as pitch or timbre) and establishing suitable thresholds for agency within the manipulation of these elements could benefit game designers and researchers alike. Lastly, the negative impact of certain sound design approaches on immersion, as revealed in the third study, creates an opportunity for a more detailed investigation. Future research could study the disruptive capabilities of sound design, exploring methods to disrupt the continuity of expression and measuring their impact on immersion. Addressing all three of these research topics would be beneficial to further reinforce the guidelines set out in the DIVE framework.

Contribution

This research has made meaningful contributions to both the field of game design research and practice, showcasing novel sound design approaches and innovative gaming mechanics within three prototypes. As this research is practice-based, the contribution encompasses both practice and research.

The development of the conceptual DIVE (Demand, Inclusivity, Versatility, Engagement) framework constitutes a significant contribution to game design research. Through interpreting the results, we determined that when designing in-game challenges, the use of sonic expression must meet the interaction expectation of the player, as stated by McMahan (McMahan, 2003). Play must be meaningful, with a discernible outcome for the player to interpret, as supported by the research of Salen & Zimmerman and expanded on by the later work of Ermi and Mäyrä. (Salen & Zimmerman, 2003) (Ermi & Mäyrä, 2005). The key findings of this research support this existing research, and this research points towards a similarity in demand that can be related back to design theory and technological affordance (Gibson, 1979) (Gaver, 1991) of devices for musical expression (Wanderley & Orio, 2002) (Tanaka, 2010).

By utilising various playing experience surveys alongside the Immersive Experience Questionnaire (IEQ) (Jennett et al., 2008), this research has yielded novel insights into the relationship between sonic, expressing, in-game challenge, and immersion. The guidelines developed within the DIVE framework aim to facilitate novel playing experiences. It is anticipated this will increase the player's level of attention, resulting in a higher level of immersion (Jennett et al., 2008) (Grodal, 2003) (Klimmt, 2003). This framework provides a theoretical foundation for understanding the role of sonic expression in immersion and the overall playing experience and will inform future researchers and game designers with a strong basis for understanding how to leverage sonic expression effectively. This is invaluable due to the current scarcity of theoretical support currently for game designs regarding the use of sound (Alves & Roque, 2011) (Collins, 2015), despite the importance of sound for affective perception (Lennart et al., 2011). The findings of this research and the DIVE framework both encourage further research in the field, challenging initial assumptions and deepening the comprehension of the role of sonic expression in immersion. Future studies utilising the DIVE framework will continue to inform and expand upon this research, ultimately enhancing our understanding of the interplay between sound, gameplay, and immersion. In conjunction, this will have practical implications for game development, as these approaches have the potential to extend beyond music games to sound design of other non-music-focused games. While the DIVE framework can be developed further with future research, it also has implications for current practice.

The prototypes developed in this research introduced novel gaming mechanics and design features. In the Preliminary Study, a music game challenged players to develop a coherent musical rhythm, shunning strict rules in favour of player expression as a core game challenge. This approach was a drastic departure from existing music games (Pichlmair & Kayali, 2007) and highlighted the potential for using sonic expression as a core gaming mechanic. The prototype in the second study required players to employ sonic expression to generate platforms and navigate a game environment in 3D. This innovative mechanic merged player-generated rhythms with procedural content generation, influencing platform properties and enemy behaviour. The success of this prototype clearly demonstrates the effective use of procedural content generation, given that, in most cases, procedural content generation often leads to less immersive level design compared to that of a human-designed level (Connor et al., 2017). The procedural content generation within the prototype for Study Two is not as complex as the generation algorithm in the work of Smith (Smith et al., 2011) or Sorenson (Sorenson et al., 2011). The dynamic between the 'low' challenge of generating individual platforms in relation to the 'high' challenge of reaching the end platform relates in an interesting (and quite literal) way to Sorenson's and Smith's 'rhythm group' approach of procedural content generation for platform and level design. (Sorenson et al., 2011) (Smith et al., 2011). This highlights the novelty of the game design and clearly demonstrates a valuable contribution to our understanding of how different uses of procedural content generation can elicit a positive impact on the player's immersion and overall playing experience. Additionally, the game featured live, runtime-generated sound design, a highly experimental approach that has not been widely seen in games and proved to be successful and well received by participants. The prototype in the third study redefined the rhythm game genre by enforcing highly precise timing to trigger non-diegetic elements accurately, with the intention of inducing input synchronisation with the remaining soundtrack elements. This was a unique approach to sound design for rhythm games and highlights potential for using in-game player interactions for controlling non-diegetic sound parameters in a novel manner.

All these contributions advance the field of game design and offer a foundation for future research and game development.

Future Directions

As outlined earlier in this chapter, the findings of this research have revealed unanswered questions that could serve as a basis for future research. The DIVE framework developed within this research is experimental, and further study in the use of sonic expression and game design would help expand the findings herein. Each of the four domains of Demand, Inclusivity, Versatility, and Engagement should be studied further, as this would deepen our understanding of the impact sound has on immersion in game design. Likewise, as discussed in the limitations of this research, it would be beneficial to iteratively study within a single gaming paradigm to provide a greater level of detail. Answering the three future research questions, as highlighted within the discussion chapter, would also prove instructive for the development of the DIVE framework:

- To what extent does prior playing experience impact the IEQ immersion factors of cognitive and challenge involvement?
- What is the relationship between dynamic range, sound elements, and immersion?
- What is the role of continuity in immersion?

Despite successfully answering the research objectives set out within this thesis, the full relationship between sonic expression and game challenge is still largely unknown. The use of sonic expression does appear to impose its own challenge, but the relationship between how this challenge interacts with aspects of a game design is still somewhat unclear. The DIVE framework aims to address this and, with further research, will be able to provide more concrete guidelines. Additionally, further work making use of the framework would work towards validating the framework.

This will be my future research direction. My immediate future practice direction will be developing the Game A prototype from the second study. The gaming mechanic for this prototype was very well received and had a lot of potential for being developed in several different directions. In the same way that rudimentary cube platforms were generated by the player's musical interactions, it is possible to generate any number of game elements. These could all be used to populate a game with rich features. This could include a narrative and (comparatively) sophisticated graphics. It would be possible to use the sonic expression to seed entire micro-game environments. These could contain a whole host of game elements from individualised regions with different physical properties, to a range of enemy compositions, to individual micro-challenges (collectables/combat/escort missions etc.). All these individual game components can be studied in relation to the use of sound, along with the relationship to the in-game challenge. This will keep me busy for the foreseeable future!

Outside of the immediate research area of game design, this research may prove useful in a computational arts or interaction design context. The more general field of HCI may find some of the concepts explored within this research intriguing and given what is known regarding sensory-modality transference within games (Lipscomb & Zehnder, 2005), it is possible that predominantly visual computing interactions may benefit from the inclusion of expressive, interactive sound. Likewise, reinforcement learning and the study of agent performance within a domain that contains expressive features may prove novel. Studying playing data and attempting to parse the motivation of sonic expression against the challenge may be an interesting task within the field of computational creativity. Lastly, an obvious choice (given the subject matter) would be gamification, specifically the gamification of music learning. It would not require a substantial amount of work to develop the prototypes to do this, as they are already relevant in their current form. Gamification in a non-musical context may also yield interesting results too, as offering sonic expression to aid visual learning may be beneficial.

Personal Reflection

Throughout my life, music has played an incredibly pivot role in forming who I am; whether acting as cultural grounding, dancing around as a small child to Motown, making friends through jamming in heavy metal bands, or as a platform for personal and professional endeavours, pursuing an undergraduate degree in Creative Music and Technology, running drum & bass and dubstep club nights, and performing either as a musician or DJ. It has always been an interest, an important facet of my life. Likewise, since the first chime of 'SEGA!' on a Mega Drive (SEGA 1988) in 1990, my world has not been the same. Games have always been my comfort zone, a place to escape to and explore, a challenge to overcome, and a test of skill. It has been deeply rewarding to combine the two, providing a wonderful platform for my imagination, along with expanding my skills and knowledge as a researcher and creative practitioner.

However, the process has been a long and difficult journey. Developing the first prototype and trying to parse my design into code was particularly hard; eventually, the process of determining a method to assess coherency was rewarding and proved deeply fascinating. The COVID-19 pandemic created a host of unknown and difficult circumstances to navigate, but I am very thankful for all the help I received from my supervisor and department. This would not have been possible without all their help. When the second study finally ran, it was exceptionally pleasing to get such meaningful results. While the prototype was planned to be considerably more complex, I am happy with how well the core challenge component worked and how well it was received by players. This was very encouraging after a particularly difficult period. To answer the key research questions and meet the objectives I had initially set out was satisfying, and it has been gratifying to deliver a framework for other game designers to use, too. As a gamer, I would be very pleased to play more games that afforded me more creative use of sound! I had originally hoped to finalise the research into a single, realised game. Unfortunately, this was not going to be possible within the timeframe and demands of research. It is fair to say I underestimated just how complex game design can get (along with all the other key game elements I didn't even focus on!). However, upon finishing this research, I will be developing the prototype from the second study into a more complete game within my own creative practice. This is an excellent outcome.

The research approach taken did have some limitations (as addressed throughout this thesis), but I am confident the approach was appropriate given my circumstances. A different approach could yield more detailed results within a single game genre but would not provide such a broad understanding. Given how sparsely populated the field of sound design with game design is, this was the correct approach. Survey design was a lot harder than I initially thought. However, once I had learned appropriate statistical analysis methods, I became confident the process was correct. This is perhaps a common theme throughout my own personal growth and development within this research. I naively thought I had a firm grasp of game design and understood the intricacies involved through my understanding as a player and as a developer of interaction sound art and very small games. I was very wrong. Games are considerably more complex than I thought, and the relationship between game design and the player is a lot more nuanced than first assumed. I have learnt a great deal both in terms of technically as a programmer/creative practitioner and as a researcher. My ability to critically reflect within both fields has vastly improved throughout this research. I suspect this relates to the largest personal challenge I faced and overcame: my innate desire to expand my knowledge and learn by taking the biggest step I can feasibly make, rather than making small, methodical steps, is not conducive to academic research. Learning to take the latter approach proved greatly beneficial for the outcome of this research, but it was not an easy one. Lastly, making novel games without an existing roadmap or model to follow was difficult and often highlighted a gap in my own technical knowledge. Learning and overcoming these gaps in knowledge was a tough challenge to overcome. Each prototype had its own complications, and I am glad to have finished each one.

Finally, I am tremendously delighted that I finished this PhD research. It has been a very long and stressful journey; I am so incredibly thankful for all the support I have received while working on it. I am excited about what further work lies ahead.

Closing Statement

There is still a great deal more to uncover regarding how sonic expression impacts the playing experience. Our knowledge of the role sound can play in immersion is still being expanded, and we have barely scratched the surface of how sonic expression can be used to make novel and engaging playing experiences. As an avid gamer, it is deeply satisfying to have contributed towards the field of game design, and as an electronic musician, I hope it encourages more practitioners to discover how much fun can be had experimenting with sound. Bleep! Bloop!

Bibliography

Accelerate, 2022-present. Accelerate Immersive Framework.

Alves, V. & Roque, L., 2009. A Proposal of Soundscape Design Guidelines for User Experience Enrichment. *Proceedings of the 4th Conference on Interaction With Sound, Audio Mostly*, pp.27–32.

Alves, V. & Roque, L., 2011. Guidelines for Sound Design in Computer Games. In *Game Sound Technology and Player Interaction: Concepts and Developments*. pp. 362–383. Available at: http://www.igi-global.com/chapter/physical-modelling-sound-synthesis/46799/ [Last Accessed 22.10.23]

Andersen F., Danny, King, C., Gunawan, A., 2021, Audio Influence on Game Atmosphere during Various Game Events, Procedia Computer Science, Vol 179, Pages 222-231, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2021.01.001.

Apple, 2002. Logic Pro.

Babakus, E., Boller, G.W., 1992. An empirical assessment of the SERVQUAL scale. In *Journal of Business Research, Vol . 24*, pp. 253-68.

Baines, J., 2015. Music 2000 Was the Greatest Way to Produce Jungle On Your Playstation. Available at: https://thump.vice.com/en_uk/article/music-2000-was-the-greatest-and-only-way-to-produce-jungle-on-your-playstation.[Last accessed 22.10.23]

Bilmes, J., 1993. *Timing Is Of The Essence: Perceptual And Computational Techniques For Representing, Learning, And Reproducing Expressive Timing In Percussive Rhythm.*MIT.

Brandse, M., Tomimatsu, K. 2013. Empirical Review of Challenge Design in Video Game Design. In: Stephanidis, C. (eds) HCI International 2013 - Posters' Extended Abstracts. HCI 2013. Communications in Computer and Information Science, vol 373.

Brockmyer, J.H. et al., 2009. Journal of Experimental Social Psychology The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4), pp.624–634.

Buttle, F. 1996. SERVQUAL: review, critique, research agenda, in *European Journal of Marketing, Vol 30 No. 1*, pp. 8-32.

Cairns, P., Cox, A. & Imran Nordin, A., 2014. Immersion in Digital Games: Review of Gaming Experience Research. In *Handbook of Digital Games*. pp. 339–361.

Calleja, G., 2011. *In-Game: From Immersion to incorporation*,

Cabello, R., 2010 – present, Three.js. Accessed at: https://threejs.org [Last accessed 09.10.23]

Carr, D. et al., 2006. *Computer Games: Text, Narrative and Play*, Available at: https://www.academia.edu/download/30721163/carr_genderplay003.pdf [Last Accessed 23.10.23]

Cassidy, G. & MacDonald, R., 2009. The effects of music choice on task performance: A study of the impact of self-selected and experimenter-selected music on driving game performance and experience. *Musicae Scientiae*, 13(2), pp.357–386.

Chang, D., Nesbitt, K.V., 2006. Developing Gestalt-based design guidelines for multi-sensory displays. In MMUI 2005: Proceedings of the 2005 NICTA-HCSNet Multimodal User Interaction Workshop, Vol 57, pp9-16.

Chion, M. & Gorbman, C., 1994. Audio-vision: Sound on Screen, Available at: https://www.jstor.org/stable/10.7312/chio18588

CodeStarrk, 2021. So many WebGL issues… Is there a stable Unity version? Every version introduces a new bug…, Unity Forums Community Discussion Platform WebGL, accessed here: https://forum.unity.com/threads/so-many-webgl-issues-is-there-a-stable-unity-version-every-version-introduces-a-new-bug.1144109/ [Last Accessed on 09.10.23]

Cohen, J., 1988. *Statistical power analysis for the behavioral sciences.*,

Collier, G. & Collier, J., 1996. Microrhythms in Jazz: A Review Of Papers. *Annual Review Of Jazz Studies*, 3, pp.117–139.

Collins, K., 2007. An Introduction to the Participatory and Non-Linear Aspects of Video Games Audio. *Essays on Sound and Vision*, (February), pp.263–298.

Collins, K., 2015. *Playing with sound*,

Conner, A.M., Greig, T.J., Kruse, J., 2017. Evaluating the Impact of Procedurally Generated Content on Game Immersion, Comput Game J 6, p209-225

Cox, A. et al., 2012. Not Doing But Thinking: The Role of Challenge in the Gaming Experience. *Proceedings of the CHI 2012*, pp.79–88. Available at: http://dl.acm.org/citation.cfm?id=2207689.

Crocker, L., & Algina, J., 1986, Introduction to classical and modern test theory, Fort Worth: Holt, Rinehart and Winston.

Cronbach, L.J., 1951, Coefficient alpha and the internal structure of tests., Psychometrika, 16(3), 297-334.

Csikszentmihalyi, 1990. Flow.

Cunningham, S., Grout, V. & Hebblewhite, R., 2006. Computer game audio: The unappreciated scholar of the half-life generation. In *Audio Mostly 2006: A Conference on Sound in Games, AM 2006 - Proceedings*. pp. 9–14.

Denisova, A. & Cairns, P., 2015. First Person vs. Third Person Perspective in Digital Games: Do Player Preferences Affect Immersion? *Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems*, 1, pp.145–148. Available at: http://dx.doi.org/10.1145/2702123.2702256.

Denisova, A., Nordin, A.I. & Cairns, P., 2016. The Convergence of Player Experience Questionnaires. *Proc. ACM CHI Play*, pp.33–37.

DeRosa, P., 2007. Tracking Player Feedback To Improve Game Design. *GameDeveloper.com*. Available at: https://www.gamedeveloper.com/design/tracking-player-feedback-to-improve-game-design [Last accessed 22.10.23]

Desain, P. & Honing, H., 1996. Physical Motion As A Metaphor For TIming In Music: The Final Ritard. In *International Computer Music Conference*. Hong Kong, pp. 458–460.

Devolver Digital, 2018. Minit.

Dobrian, C. & Koppelman, D., 2006. The E in NIME: Musical Expression with New Computer Interfaces. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp.277–282.

Double Fine Productions, 2019, Brutal legend (PS3, Xbox 360, PC, Mac, Linux).

Drool, 2016. Thumper.

Eidos, 2000. Hitman.

Ermi, L. & Mäyrä, F., 2005. Fundamental Components of the Gameplay Experience: Analysing Immersion. *Changing Views: Worlds in Play*, pp.15–27.

Fake, N., 2004. The Sky Was Pink [06BC], UK, Border Community.

Fake, N., 2005. Issst #2 [FOR 1100 6], Germany, Four Music.

Fels, S., Gadd, A. & Mulder, A., 2002. Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, 7(2), pp.109–126.

Fencott, P., 1999. Presence and the Content of Virtual Environments. Available at: https://www.fencott.com/Clive/TU/research/Fencott1999Content.pdf [Last accessed 22.10.23]

Field, A., Miles, J. & Field, Z., 2013. *Discovering Statistics Using SPSS*,

Fitterer, D., 2008. Audiosurf (PC).

Friberg, A. & Sundberg, J., 1999. Does Music Performance Allude To Locomotion? A Model Of Final Ritardandi Derived From Measurements Of Stopping Runners. *The Acoustical Society of America*, 105, pp.1469–1484.

Frost, J., How to Correctly Interpret P Values Title.

Gaver, W.W., 1991. Technology Affordances. , pp.79–84.

Gibson, J.J., 1979. The Ecological Approach to Visual Perception. *Houghton Mifflin-Boston*.

Goldsmiths University of London, 2023. Data Protection available at: https://www.gold.ac.uk/data-protection/ [Last accessed 18.10.23]

Grierson, M., 2009. Maximilian.

Grimshaw, M., 2008. Sound and immersion in the first-person shooter. *International Journal of Intelligent Games & …*, pp.119–124. Available at: http://wlv.openrepository.com/wlv/handle/2436/35995. [Last accessed 22.10.23]

Grimshaw, M., Charlton, J.P. & Jagger, R., 2011. First-Person Shooters: Immersion and Attention. *Eludamos. Journal for Computer Game Culture*, 5(1), pp.29–44. Available at: https://eludamos.org/index.php/eludamos/article/view/vol5no1-3 [Last accessed 22.10.23]

Grimshaw, M. 2014. Sound in: The Routledge Companion to Video Game Studies, edited by Wolf, P.J.,Perron, B., p310-326

Grodal, T., 2003. Stories for Eye, Ear, and Muscles Computer Games, Media, and Embodied Experiences. *The video game theory reader*, pp.129–155. Available at: https://www.taylorfrancis.com/chapters/edit/10.4324/9780203700457-7/stories-eye-ear-muscles-torben-grodal [Last accessed 22.10.23]

Harmonix, 2005. Guitar-Hero (PS2).

Harmonix, 2016. Music VR (PS4).

Harmonix, 2008. Rock-Band (Multi-Platform).

Harvey, A., 2009. Seeking the Embodied Mind in Video Game Theory: Embodiment in Cybernetics, Flow, and Rule Structures. *Loading…Journal of the Canadian Game Studies Association*, 3(4), p.9.

Hello Games, 2015. No Man's Sky. (Switch, PS5, PS4, Xbox One, macOS, Windows)

Hendrikx, M, Meijer, S, Van Der Velen, J, IOSUP, A., 2012. Procedural Content Generation for Games: A Survey. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2012–Jan(January), p.24.

Hilbert, D.M. & Redmiles, D.F., 2000. Extracting usability information from user interface events. *ACM Computing Surveys*, 32(4), pp.384–421.

Hills, J., 2022. The role pressure plays in game design. Available at: https://joshhills.dev/article/pressure-in-game-design [Last Accessed 09.10.23]

Hinton, P R., Brownlow, C., McMurray, I., Cozens, B., 2004, SPSS Explained, Routledge, p 364.

Holden, J., 2003 – present. Border Community.

Holden. J., 2004. The Sky Was Pink [06BC], UK, Border Community.

Hook, A., & Coulton, P. (2017). Games Design Research through Game Design Practice. In P. Lankoski, & J. Holopainen (Eds.), Game Design Research: An Introduction to Theory and Practice (pp. 97). ETC Press.

Hsu, S.H., Wen, M.H., Wu, M.C. (2007) Exploring design features for enhancing players challenge in strategy games. Cyberpsychology & Behaviour 10 (3), 393-397

Huiberts, S., 2010. Captivating Sound: The Role of Audio for Immersion in Computer Games., (November), pp.1–200.

Huiberts, S. & van Tol, Ri., 2008. IEZA: A Framework for Game Audio. *GameDeveloper.com*, Available at: https://www.gamedeveloper.com/audio/ieza-a-framework-for-game-audio [Last accessed 22.10.23]

Huizinga, J., 1968. Homo Ludens: a Study of the Play Element in Culture.

Huron, D., 2008. Lost in music. In Nature 453, pp 456-457. Available at: https://www.nature.com/articles/453456a#citeas [Last Accessed 09.10.23]

IJsselsteijn, W. a., de Kort, Y. a. W. & Poels, K., 2013. Game Experience Questionnaire. , pp.1–9.

Indieszero, 2005. Electroplankton (DS).

Iyer, V., 2002. Embodied Mind, Situated Cognition, and Expressive Microtiming in African-American Music. *Music Perception: An Interdisciplinary Journal*, 19(3), pp.387–414.

Jamieson, D., 2017. The Elements of Game Design: Part 1. Available at: https://code.tutsplus.com/four-elements-of-game-design-1--cms-22720a [Last Accessed 08.10.23]

Jennett, C. et al., 2008. Measuring and defining the experience of immersion in games. *International Journal of Human Computer Studies*, 66(9), pp.641–661.

Jennett, C.I., 2010. Is game immersion just another form of selective attention? An empirical investigation of real world dissociation in computer game immersion. , (July), p.243.

Jester, 1999. Music 2000 (PS).

Juul, J., 2003. The Game, The Player, The World: Looking for a Heart of Gameness. In: Copier, M., Raessens, J. (eds.) *Level Up: Digital Games Research Conference Proceedings*, pp.30–45.

Juul, J., 2009. Fear of Failing? The Many Meanings of Difficult in Video Games, edited by Wolf, P. J., Perron, B., in The Video Game Theory Reader 2.

Kearney, P.R., 2007. Immersed and how? That is the question. *Game in' Action*, p.12. Available at: https://www.researchgate.net/publication/255651360_Immersed_and_how_That_is_the_question [Last accessed 22.10.23]

Keil, C. & Feld, S., 1994. *Music Grooves*, Chicago: University of Chicago Press.

Kenwright, B., 2020, 'There's more to Sound than Meets the Ear: Sound in Interactive Environments', IEEE Computer Graphics and Applications, vol. 40, no. 4, pp. 62-70., http://doi.org/10.1109/MCG.2020.2996371

Kim, J.H. et al., 2008. Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems. *CHI '08 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp.443–451. Available at: http://dl.acm.org/citation.cfm?id=1357126. [Last accessed 04.10.18]

Kim, S., 2004. What is a Puzzle?

King, 2012. Candy Crush Saga (iOS, Android).

Klarlund, M., Brattico, E., Pearce, M., Wu, Y., Vuust, P., Overgaard, M., 2023. Words apart? Testing the cultural distance hypothesis in music perception of Chinese and Western Listeners. In Cognition Vol. 235.

Klimmt, C., 2003. Dimensions and Determinants of the Enjoyment of Playing Digital Games : A Three-Level Model. In *Level Up: Digital Games Research Conference*. pp. 246–257.

Koçak, C., Egrioglu, E., Yolcu, U., & Aladag, C., 2014, Computing Cronbach Alpha Reliability Coefficient for Fuzzy Survey Data. American Journal of Intelligent Systems. 4. pp. 204-213. 10.5923/j.ajis.20140405.03.

Konami, 1998. Metal Gear Solid (PS).

Korg, 2003. Korg Electribe ER-1.

Kurtz, G., 1998. Games and Instruments: Two Ways to Play.

Leman, M., 2007. Embodied Music Cognition and Mediation Technology. *Processing*,

Lennart, E.N., Grimshaw & Marc, 2011. Player-Game Interaction Through Affective Sound. *Game Sound Technology and Player Interaction: Concepts and Developments*, 51, pp.264–285.

Lipscomb, S.D. & Zehnder, S.M., 2005. Immersion in the virtual environment: the effect of a musical score on the video gaming experience. *Journal of Physiological Anthropology and Applied Human Science*, 23(6), pp.337–343.

Lobanov, G. 2018, Wandersong (Mac, PC, Switch, PS4, Xbox One).

Löfgren, K., Normality test using SPSS: How to check whether data are normally distributed. *Youtube.com*. Available at: https://www.youtube.com/watch?v=IiedOyglLn0 [Last accessed 22.10.23].

Ludopium, 2019. Vectronom (Switch, Android, iOS, PC, Mac).

Marcos, D., McCurdy, D., Ngo, K., 2015. A-frame. Available here: https://github.com/aframevr/aframe [Last Accessed: 09.10.23]

McMahan, A., 2003. Immersion, Engagement, and Presence: A Method for Analyzing 3-D Video Games. *The Video Game Theory Reader*, pp.67–86.

Microsoft, 1997-present. Visual Studio.

Mimic Project, 2019. MaxiInstruments.

Monstars, 2015. Rez Infinite (PS4).

Nacke, L. et al., 2011. Biofeedback game design: using direct and indirect physiological control to enhance game interaction. *Proceedings of the SIGCHI …*, pp.103–112. Available at: http://dl.acm.org/citation.cfm?id=1978958. [Last accessed 22.10.23]

Nacke, L., Lindley, C. & Stellmach, S., 2008. Log Who's Playing: Psycholophysiological Game Analysis Made Easy Through Event Loggin. *International Conference on Fun and Games, Eindhoven*, pp.150–157.

Nacke, L.E. et al., 2009. Playability and Player Experience Research. *Proceedings of the IEEE*, pp.1–11.

NanaOn-Sha, 1996. Parappa the Rapper (PS).

NanaOn-Sha, 1999. Vib-Ribbon (PS).

Neustaedter, C., & Sengers, P., 2012. Autobiographical design: What you can learn from designing for yourself. interactions. 19. pp.28-33.

Ng, P., & Nesbitt, K., 2013, Informative Sound Design in Video Games. ACM International Conference Proceeding Series. 10.1145/2513002.2513015.

Nintendo, 1998. The Legend of Zelda: Ocarina of Time (N64).

Nintendo, 2000. The Legend of Zelda: Majora's Mask (N64).

Nordin, A. Imran, 2014, Immersion and Players' Time Perception in Digital Games.

Nordin, A. Imran, Denisova, A., and Cairns, P., 2014, 'Too Many Questionnaires: Measuring Player Experience Whilst Playing Digital Games', Conference: The Seventh York Doctoral Symposium on Computer Science and Electronics.

Nordkvist, P., 2006. Some Polyphony [12BC], UK, Border Community.

Nutting Associates, 1971. Computer Space (Arcade).

Oculus VR, 2014 - 2021. Oculus Rift Development Kit 2.

Orvis, K.A., Horn, D.B., Belanich, J., (2008) The roles of task difficulty and prior videogame experience on performance and motivation in instructional videogames. Computers in Human Behaviour 24 (5), 2415-2433

Pagulayan, R. & Steury, K., 2004. Beyond Usability In Games. *Interactions*, 11, pp.70–71.

Pagulayan, R.J. et al., 2012. User-centered Design in Games. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, pp.883–906.

Pichlmair, M. & Kayali, F., 2007. Levels of Sound: On the Principles of Interactivity in Music Video Games. *DiGRA '07 - Proceedings of the 2007 DiGRA International Conference: Situated Play*, 4, pp.424–430. Available at: http://www.digra.org/dl/db/07311.14286.pdf. [Last accessed 22.10.23]

Plummer, K., Reporting The Wilcoxon Signed-Ranks Test. Available at: https://www.slideshare.net/plummer48/reporting-the-wilcoxon-signed-ranks-test [Last accessed December 1, 2017].

Poepel, C., 2005. On Interface Expressivity: A Player-Based Study. *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp.228–231. Available at: http://www.nime.org/proceedings/2005/nime2005_228.pdf [Last accessed 22.10.23]

Reality Labs, 2020 – present. Quest 2.

Reiner, M. & Hecht, D., 2009. Behavioral indications of object-presence in haptic virtual environments. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, 12(2), pp.183–186.

Repp, B., 1990. Patterns Of Expressive Timing In Performances Of A Beethoven Minuet By Nineteen Famous Pianists. *The Acoustical Society of America*, 88, pp.622–641.

Rigby, S. & Ryan, R., 2007. The player experience of need satisfaction (PENS) model. *Immersyve Inc*.

Risi, S. et al., 2014. Automatically Categorizing Procedurally Generated Content for Collecting Games.

Roland, 1983 - 1985. Roland TR-909.

Russell, D., 2012. Video Game Audio: Diegesis Theory. *Dev.Mag*. Available at: http://devmag.org.za/2012/04/19/video-game-audio-diegesis-theory-2/ [last accessed 22.10.23]

Salen, K. & Zimmerman, E., 2003. Rules of Play: Game Design Fundamentals. *MIT Press*, 2004, p.688.

Sanders, T., & Cairns, P., “Time perception, immersion and music in videogames,” in Proceedings of the 24th BCS Interaction Specialist Group Conference. British Computer Society, 2010, pp. 160–167

Samuels, P., Marshall, E. & Lahmar, J., 2012. Checking normality for parametric tests in SPSS, pp.1–4. Available at: https://www.researchgate.net/publication/320831642_Checking_Normality_for_Parametric_Tests [Last accessed 22.10.23]

Sauro, J.,2014, What Does Statistically Significant Mean? *MeasuringU*. [Last accessed 04.10.18]

Shapiro, S.S. & Wilk, M.B., 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), p.591. Available at: http://www.jstor.org/stable/2333709?origin=crossref [Last accessed 22.10.23]

Schaffhausen, A., Raabe, W., 2006. Schmedding 8000 [TRAUM V79], Germany, Traum Schallplatten.

Schmalbach, J., 2005. Issst #2 [FOR 9041 6], Germany, Four Music.

Siroker, D., Koomen, P., 2015, A / B Testing: The Most Powerful Way to Turn Clicks Into Customers, p8,

Slater, M. et al., 2009. How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicologia*, 40(2), pp.193–210.

Slater, M., 1999. Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 8(5), pp.560–565. Available at: http://www.mitpressjournals.org/doi/10.1162/105474699566477. [Last accessed [22.10.23]

Slater, M., 2002. Presence and The Sixth Sense. *Presence: Teleoperators and Virtual Environments*, 11(4), pp.435–439. Available at: http://www.mitpressjournals.org/doi/abs/10.1162/105474602760204327. [Last accessed 22.10.23]

Smith, G. et al., 2011. Launchpad: A rhythm-based level generator for 2-D platformers. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(1), pp.1–16.

SomethingElse, 2013. Papa Sangre.

Sorenson, N., Pasquier, P. & DiPaola, S., 2011. A generic approach to challenge modeling for the procedural creation of video game levels. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(3), pp.229–244.

Square Enix, 2015. Rise of The Tomb Raider (PS4, Xbox 360, Xbox One, Windows)

Steinberg, 1989. Cubase.

Sweetser, P. & Wyeth, P., 2005. GameFlow: A Model for Evaluating Player Enjoyment in Games. *Comput. Entertain.*, 3(3), pp.3–3. Available at: https://dl.acm.org/doi/10.1145/1077246.1077253 [Last accessed 22.10.23]

Tafalla, R.J., 2007. Gender differences in cardiovascular reactivity and game performance related to sensory modality in violent video game play. *Journal of Applied Social Psychology*, 37(9), pp.2008–2023.

Tan, S.-L., 2014. Video Games: Do You Play Better With the Sound On or Off?

Tanaka, A., 2010. Mapping Out Instruments, Affordances, and Mobiles. *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, (Nime), pp.15–18. Available at: http://www.nime.org/proceedings/2010/nime2010_088.pdf. [Last accessed 22.10.23]

Tavakol, M. & Dennick, R., 2011. Making sense of Cronbach ' s alpha. *International Journal of Medical Education*, pp.53–55.

Thaut, M.H., Trimarchi P.D., Parsons, L.M., 2014. Human Brain Basis of Musical Rhythm Perception: Common and Distinct Neural Substrates for Meter, Tempo, and Pattern. Brain Sci p428-452.

Thompson, M., Nordin, A. Imran., and Cairns, P. (2012). Effect of Touch–Screen Size on Game Immersion. 280-285. 10.14236/ewic/HCI2012.38.

Thubron, R. 2020. The Future of Tech: Gaming Consoles, the Xbox and Playstation of Tomorrow, *techspot.com* available at: https://www.techspot.com/article/1979-the-future-of-game-consoles/ [Last accessed 07.10.23].

Tobar, R., 2008. With You [TRAUM V103], Germany, Traum Schallplatten.

Todd, N., 1989. Towards A Cognitive Theory Of Expression: The Performance and Perception of Rubato. *Contemporary Music Review*, 4, pp.405–416.

Tri-Crescendo, 2007, Eternal Sonata (Xbox 360)

Tychsen, A. & Canossa, A., 2008. Defining personas in games using metrics. *Proceedings of the 2008 Conference on Future Play Research, Play, Share - Future Play '08*, pp.73–80. Available at: http://portal.acm.org/citation.cfm?doid=1496984.1496997 [Last accessed 22.10.23]

Ubisoft, 2002. Tom Clancy's Splinter Cell.

UnitedGameArtists, 2001. Rez (Dreamcast).

Unity Technologies, 2004-present. Unity.

Virtuix, 2017. Viruix Omni.

Wanderley, M.M. & Orio, N., 2002. Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI. *Computer Music Journal*, 26(3), pp.62–76.

WARP, 1997. Real Sound: Kaze no Regret.

Wesley, D. & Barczak, G., 2010. Innovation and Marketing in the Video Game Industry: Avoiding the Performance Trap. In *Innovation & Marketing in the Video Game Industry: Avoiding the Performance Trap*. pp. 170–194.

Wiseheart, M., Effect Size Calculator. *webpage*.

Wöhrman, S., & Ningalei, N., 2018. The Impact of Sound on Player Experience - A literature study on how players experience the encounter with sound in horror-games. Available at: https://www.diva-portal.org/smash/get/diva2:1480010/FULLTEXT01.pdf [Last accessed 15/05/24].

Yamada, M., 2002. I. In *7th International Conference on Music Perception and Cognition*. pp. 340–343.

Yamada, M., Fujisawa, N. & Komori, S., 2001. The Effect Of Music On The Performance And Impression In A Video Racing Game. *Journal of Music Perception and Cognition*, 7(2), pp.65–76.

Yee-King, M., & Dall'Avanzi, I., 2018. Procedural Audio in Video Games. In: Newton Lee, ed. Encyclopedia of Computer Graphics and Games. Cham, Switzerland: Springer. ISBN 9783319082349

Appendix

Preliminary Study: Game Design Document

Puzzle Prototype

Working Title: Boxel

Introduction

This puzzle game introduces the player to broad concepts of musical rhythm development, and to simple musical writing concepts such as rotation and inversion of rhythms. This is with the intended aim to create a novel playing experience using the player's input to create a self-impose game challenge.

Gameplay

The player is instructed to build a rhythm. They are given four repetitions of a four-bar loop to add to their rhythm, with each successive bar-input being checked against the previous bar; comparing the distribution of notes per subdivision per bar as it moves though the loop. At the end of the four-bar loop the salient features of the loop are used to generate the backing track for the proceeding four-bar rhythm and the user starts again on a new level with a blank four-bars to fill.

To make progress within the game, the player must continually develop coherent rhythmic structures based on their previous input. For simplicities sake, we are defining coherency to be pattern matching of small rhythm fragments from the previous rhythm to the newly created one, the similarity of note distribution between subdivisions relative to the beat, the bar, and the four-bar loop, and how these note distributions imply or characterise the metric pulse (either 'pushing', 'pulling', or explicitly stating the emphasis of the beat).

If there is not sufficient input per bar or per repetition the player's rewards diminish, ensuring that players are rewarded for novel interactions. If the player's input stagnates over two of the four-bar loops the player loses the game. Additionally, if the player's input is not coherent across the four-bar loop twice in a row the player also loses the game.

Game Environment

The game environment shares similarities with a traditional step-sequence synthesiser. However, rather than displaying each step of the sequence at once with a moving 'play-head' the game environment uses a fixed camera position acting as play-head and instead moves the steps based on their position relative to the play-head accordingly. This allows the player to make cursory glances at incoming structures whilst giving them a consistent visual location of where the play-head is (at the bottom of the screen).

The sequencer also exists in 3D with the different sounds arranged across the x and z axes evenly. As the player activates a step in the sequencer, they are simultaneously building simple 3D structures comprised on these cubes. This provides visual feedback and reinforce the player's understanding of their generated rhythmic structures. As well as, visually highlighting where points of rotation or inversions can occur - as transformations based on symmetry such as rotations or inversions are easily communicated visually, by simply moving these representations around the centre point of the x, z axes.

Rhythm Checking

The pipeline would be as follows:

1. The player adds a note to the sequence
2. We check the note position relative to the backing track
3. We check note position against the previous beat
4. We check against previous bar
    a. Are there (comparatively) any rotations?

    b. Are there (comparatively) any inversions?

5. We check entire distribution across the bar against previous bar distribution
6. We check symmetry of individual drum notes and accents across two bars
7. We check symmetry of individual drum notes and accents across the full loop
8. End of pipeline, wait for step 1.

Levels

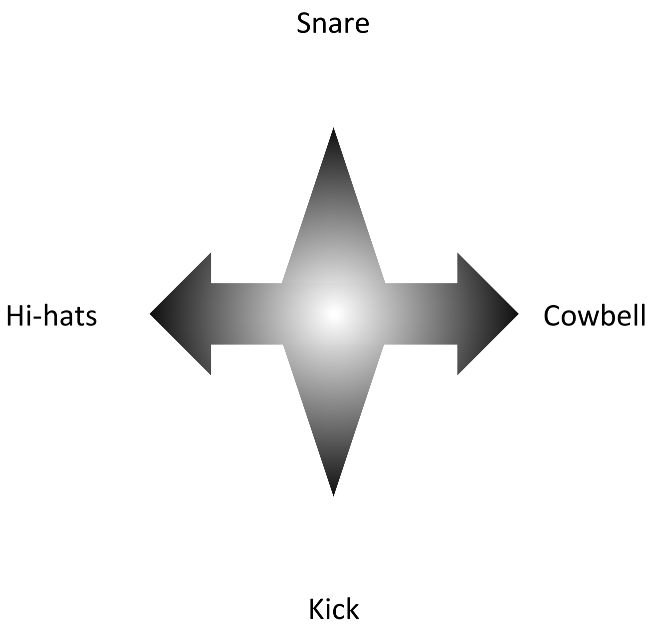
Although the task of each level is the same, each level will increase the tempo of the loop (starting at 118 bpm), meaning players will have to respond quicker every cycle. As each rhythm is checked against the previous rhythms, the player cannot use the same structures continually, requiring each successive input to deviate from the previous, therefore creating a self-imposed challenge.

Scoring

Players accumulate points based on how the construct their rhythm, scoring multipliers for the use rotations of smaller sub-rhythms within a bar, inversions across loops, and for consecutive coherent placement of notes. If the player fails to place notes coherent their score loses points, if their input stagnates their score starts to diminish until their input shows deviation again. If the player fails to develop a rhythm after two full cycles the game is over.

Controls

The player has a choice of four different percussive sounds to trigger. these are chosen by swiping the touch screen in a direction: up, down, left, or right. The length of the swipe generates the notes velocity allowing the player to add accents to their rhythms. Currently, the four sounds are setup as follows:



Each axis positions the typically-opposing accenting percussive sound in the opposite direction to allow the player to make easier and faster decisions based on intuition.

Interface

There will be a short tutorial for the first time playing in which the players are introduced to the controls, the game environment, and basic concepts such as symmetry, rotation, and inversion. The only explicit feedback on screen will the score and current level/tempo they are currently on.

If the player's input stagnates the game-space will demonstrate an example of a rotation or inversion by displaying the corresponding shape comprised of the sequencer cubes in the form of a superimposed semi-translucent cubes.

Art

The art design is clean and concise, using simple bold colours to differentiate between cube pieces and background. Each step of the sequence is single cube, displayed on a two-by-two cube grid on the x, z axis. These are placed inside a larger semi-transparent two by two sized cube to highlight that each cube is itself a subdivision of a division of a beat and bar. With four of these larger divisions displayed on the screen at any one time. The incoming steps are displayed on the y-axis above the corresponding coordinate for the note. At the end of each completed level the player's complete rhythm structure is highlighted within the four larger semi-transparent cubes. Before fading out as a new level starts.

Additionally, the background will be a two-tone colour gradient moving from the incoming level at the top of the screen to the current level background colour at the bottom, as the player moves through each repetition of loop the gradient moves lower until the screen is filled with the new colour. Implying the player's position in the level.

Sound

The four percussive sounds will be samples from a Roland 909. The backing track will be also comprised of electronically generated samples, with the arrangement being based on the player's previous rhythmic input. This will be achieved with a simple ruleset based on distribution of notes and of accents within the previously generated level. As the backing track is sample based it allows for quick changing of sounds, meaning each level could have its own unique sound in terms of style and timbre. However, these will be limited to sounds clearly electronic in origin – serving both as a choice of aesthetics and to embrace the origins of the gaming mechanism – the step sequencer, and to hide the limitations the sequencing mechanism currently has – such as not being able to account for micro-rhythms of steps, swing, triplets or other non-even note divisions.

Preliminary Study Introductory Video Script

(Taken from the video)

Welcome to Boxel, a rhythm- based game.

To start swipe anywhere on the touch screen.

The aim of the game is to score points by building a coherent rhythm.

If you swipe in a direction you generate a rhythm piece. Down for a kick, left for hi-hat, up for a snare, right for a cowbell.

When you swipe and generate a sound they will appear at the top of the tower before moving down.

Each cube piece is a 16^{th} musical note, if you don't know what that means, that's ok.

The idea is to simple swipe and generate these cubes into a pattern.

The pattern loops after two bars of music, or 32 of those individual cube pieces.

As the loop plays through, the cube pieces move towards the bottom of the screen.

You are awarded points for how well you match and deviate from your initial pattern.

If you deviate too far, you will score negative points. This is also the same for the lack of interaction.

Your score is displayed at the bottom of the screen.

You will also find a bar counter at the top of the screen.

Each level has 32 bars in it.

Once you reach the end of the level, your pattern will slowly be removed, along with cube pieces.

First the kicks, then the hi-hats, then cowbells, then finally the snares.

This allows you more spaces to continue building your pattern.

Preliminary Pre-Survey Example

1. **Name:**

(for referencing/matching questionnaires only, will be kept off record If a service like survey monkey is used will need to match surveys together)

2. **Age:**

   18-29

   30-39

   40-49

   50-59

   60+

3. **Do you play video-games?**

YES – **go to question 5**

NO

4. **Would you ever consider playing video-games?**

YES

NO – **end of survey you are not a suitable candidate for this study**

5. **How Long have you played video-games for?**

Less than 1 week

Less than 1 month

Less than 1 year

Less than 5 years

5 years +

6. **How frequently do you play?**

A couple times per year

A couple times per month

About one day a week

2-4 days per week

5 or more days per week

7. **What devices do you usually play on? Tick multiple where applicable.**

Phone/Tablet

Nintendo DS or other handheld game specific device

Console with standard non-motion-based controller such as PS or Xbox

Console with motion-based controller such as Nintendo Wii or Xbox WITH a Kinect available

PC with standard non-motion controller

PC with motion controller

VR, please name which:

If answered any motion-based move to question 8 else jump to question 9.

8. **How frequently do you play games that utilise player movement?**

A couple times per year

A couple times per month

About 1 time per week

2-4 times a week

5 or more times per week

9. **What types of games do you usually play? Tick multiple where applicable.**

Action/Adventure

Casual

Fighting

First-Person-Shooter

Racing

Simulation

Sports

Strategy

Other, please name:

10. **Which of the following music games have you played? Tick multiple where applicable.**

Guitar Hero

Music VR

Rez

Rock Band

Thumper

Other, please name:

None

11. **Do you have any experience playing a musical instrument?**

YES, state which:

NO – **Jump to question 14**

12. **How many years have you played?**

Less than 1

2-4 years

5-6 years

7+ years

13. **How often do you practice?**

A couple times per year

A couple times per month

About once a week

2-4 times per week

5 or more times per week

14. **Do you have any experience using music production software such as Ableton Live or Cubase?**

    YES

    NO – **Jump to question 17**

15. Do you still make music/currently a digital music practitioner?

YES – **jump to question 17**

NO

16. Why did you stop?

Lack of interest

Not enjoyable

Not sociable

Time constraints

Too difficult to improve

Other, please name:

17. How were you first introduced to musical concepts?

1 on 1 instrument lessons with a teacher/tutor

Group lessons in a class with a teacher

At home with parent/guardian

Self-Taught

Other, please name:

Never been introduced to musical concepts.

Preliminary Post-Survey Example

Please select your choice

1. **Name:**

(for referencing/matching questionnaires only, will be kept off record If a service like survey monkey is used will need to match surveys together)

2. **How did you enjoy your playing experience?**

Not at all				A lot
1	2	3	4	5

3. **How clear was the game objective?**

Not at all				Very
1	2	3	4	5

4. **How challenging was the game?**

Not at all				Very
1	2	3	4	5

5. **How did you feel in control of the in-game challenge?**

Not at all				Very
1	2	3	4	5

6. **How clearly did you understand the effect your note placement had on your score?**

Not at all				Very
1	2	3	4	5

7. How intuitive was the interaction?

Not at all				Very
1	2	3	4	5

8. How responsive was the input on the sound in the game?

Not at all				Very
1	2	3	4	5

9. How aware were you of the changes in the accompanying backing track?

Not at all				Very
1	2	3	4	5

10. What was the most important feedback that helped?

By ear

By visual cues

11. How useful did you find the visual feedback of the game useful?

Not at all				Very
1	2	3	4	5

12. **Did you understand where your position in the musical loop was?**

Never				Always
1	2	3	4	5

Regarding the visual feedback and your position in the musical loop:

13. **How important was the bar counter?**

Not at all				Very
1	2	3	4	5

14. **How important were the moving semi-transparent cubes?**

Not at all				Very
1	2	3	4	5

15. **How important that the cubes opacity increased as the loop reached its starting point?**

Not at all				Very
1	2	3	4	5

16. **How much did it feel like a musical experience?**

Not at all				Very
1	2	3	4	5

17. How likely would you play this game again?

Not at all				Very
1	2	3	4	5

Study Two Game Design Document

Platformer Prototype

Working Title: Red is Dead

Introduction

This 3D VR platform game uses musical rhythm as a gaming mechanic to dynamically alter the game environment; generating platforms for the player to utilise in order to navigate to an 'end goal' platform. It also generates enemies that will obstruct the player, which need to be avoided. The intended aim is to create a novel playing experience using the player's input as musical expression, coupling the self-imposed challenge of rhythm generation and the predetermined in-game challenge.

Gameplay

The player is instructed to reach an end goal platform, visible to the player by looking up. To do so they are instructed to collect tokens as they appear in the level on top of the platforms in the game environment. Each platform has a red enemy 'playhead' that moves along each row of user generated platforms. The players soon discover these are to be avoided as they push the player off the platforms. If the player gets knocked off or falls off, they will fall until they hit a red kill plane. This then places them back on the initial starting platform. Each row of platforms is comprised of 16 tiles. Each tile corresponds with a 16^{th} note in a step sequencer that has a total of 16 steps. The enemy playheads position corresponding with the master clock. It is the current position of global playhead (the point in which we are trigger sound / steps in our audio loop). In version A once players collect a token, they can press one of four buttons to place a sample in the step of the sequencer they are standing on. Each button corresponds with a different percussive sound. The player also has a button to delete a placed token. Once a token is place, in-conjunction with the sample being placed in that step, a new row of platforms is created 1 unit up on the y-axis. In version B of the game, players do not have any say over what sounds they place or where they are placed. Upon pickup of the token a sample is dropped in that position of the step sequencer. This now plays the sample in that position of the loop every time the global playhead reaches that point. The players other interactions also impact the sonic landscape of the game. Jumping, falling, and contact with the playheads also change parameters of the sonic environment. By collecting tokens, avoiding the playhead enemies, and generating and climbing the platforms, players are able to navigate to the end platform.

Game Environment

The game environment always has a white static starting platform. This is where the player starts and is returned if they fall. The game always a black end goal platform. This is the target platform the player is trying to get to in order to win the game. This platform follows the players x and z position with some easing. Consequently, regardless of which direction the player builds their platforms around the environment, the end goal is always reachable. The game environment also always starts with one row of platforms in front of the player. As discussed above, these are comprised of 16 tiles, each corresponding with a step in a step sequencer. These platforms come in four different colour gradients, depending on which sample is used, each with their own corresponding shape. The initial row of platforms is a gradient of white to black. When a new platform is generated, it starts from the current tile the player is standing on, although its global step position is rotated by that current step value (see explanation under the heading 'Platform Generation' below for more detail). For every platform there is a new token to collect and a corresponding playhead to avoid. The shape of playheads also corresponds to the platform / sample type used. As the player collects more tokens the backing music changes over time, moving through different arrangements.

Platform Generation

When a player has a token and presses a button to place a sound and generate a new row of platforms, it does so from the current position / tile the player is standing on. It generates the new platform 1 unit on the y-axis above. The new platform will be positioned in front of the player. However, the first tile in the row of new platforms does not correspond with the first step of the sequencer. It corresponds with previous step the player was on; all new tiles are rotated around this step value. For example:

```
PlayerPosition:                P
Platform 1 / Tiles:  0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Playhead 1:          E
```

Player uses the token on tile 5 of platform 1.

The game now has the following:

```
PlayerPosition:                P
Platform 1 / Tiles:  0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Playhead 1:          E

PlayerPosition:      P
Platform 2 / Tiles:  5 6 7 8 9 10 11 12 13 14 15 0 1 2 3 4
Playhead 2:                                      E
```

Once the global playhead trigger moves a single unit both enemy playheads will be positioned above tile 15. They cycle sequentially forwards. If the player collects another token and moves along the new platform and places it on the 9th tile, we would get the additional platforms:

```
PlayerPosition:              P
Platform 2 / Tiles:  5 6 7 8 9 10 11 12 13 14 15 0 1 2 3 4
Playhead 2:                                      E
PlayerPosition:      P
Platform 2 / Tiles:  9 10 11 12 13 14 15 0 1 2 3 4 5 6 7 8
Playhead 2:                              E
```

Along with the original platform and playhead. Each sample and corresponding platform type store which tiles have been used for that sound. Subsequently, for each new platform of same type we can get that placed sample on the new platform. This allows the player to build interesting structures they can use to help them navigate around the moving playheads.

Rhythm Checking

The input rhythm of the player's sequence has its distribution checked and scored. If a player uses too many of certain sounds in a distribution deemed incoherent, it has an impact on where the following token is placed.

Levels

There is a single level with a defined end. Once players complete the level, they are given 4 bars to listen to their rhythm before each section is removed. The game environment assets are sequentially removed or shaded black. After 4 bars the players are taken to a never-ending bonus level where they can retrigger samples in place, without the end for tokens.

Winning

In order to win players must reach the end platform. However, there is no time limit on how long it takes to do so.

Controls

The facing direction of the players head is used for the camera's look control. Players can move their heads and look around the 3D environment. They can use the analogue sticks on either controller to navigate around the game space. Both can be used at the same time allowing the player to turn in an arc. The forward direction is dependent on their facing direction. This makes it easier / more intuitive for players to move. The four triggers on the Oculus touch controllers are used for the different sounds, with a button on each controller for jumping and for destroying tokens in the Game A version. In version B you only have the jump and movement to control.

Thumbsticks + facing direction = forward

Thumbsticks = strafing

Left trigger = kick drum

Right trigger = snare drum

Left grip trigger = hi-hat bleep

Right grip trigger = bell bloop

X or A = Jump

Y or B = destroy steps you're currently on.

Interface

The game is presented in an A-frame canvas on a webpage. The webpage has the control guide and the game information on how to play. They are instructed to enter VR mode to start. The game does not work unless on an Oculus Quest device.

Art

The art design is deliberately clean and concise. It uses simple bold colours to differentiate between the different platform types, the collectable tokens, and has a clear message that red objects are to be avoided. The playheads have some transparency so players can see where the tokens are through them. It is deliberately minimal to not distract from the sound in the game. What little choice is visual aesthetic (such as colour palette) was made in homage to early PC games of the 80s and some early 90s.

Sound

The four percussive sounds are sampled from various groove boxes. They are electronically generated samples, deliberately chosen for their suitability for the desired retro inspired glitch / IDM style. This style was chosen as it lends itself to rapid changes if applied effects. As mentioned briefly in the gameplay topic above, player movement and interaction changes parameters of the backing synths. Player movement such as jumping also changes sample rate playback which adds interesting timbral changes when applied ad hoc to our percussive samples. This is a typical feature of our chosen style. Additionally, there is a bass synth with parameterised LFO, filterFrequency, PitchMOD, a midrange backing synth with parameterised features, and a synth for collision of our enemy playheads, which trigger a note from an array. These are subject to audio effects such as delay and reverb, which are also parameterised based on in-game activity. In the B version of the game, the rhythmic sequence (and subsequently platform structure) has been designed to create an end result rhythm that is coherent to the style of the chosen electronic music genre.

Study Two Survey Example

1. Name:

(for referencing/matching questionnaires only, will be kept off record If a service like survey monkey is used will need to match surveys together)

Demographic Questions

2. Age:

18-29, 30-39, 40-49, 50-59, 60+

3. How Long have you played video games for?

Less than 1 week, Less than 1 month, Less than 1 year, Less than 5 years, 5 years +

4, **How frequently do you play?**

A couple times per year

A couple times per month A

bout one day a week

2-4 days per week

5 or more days per week

5. What devices do you usually play on? Tick multiple where applicable.

Phone/Tablet

Nintendo DS or other handheld game specific device

Console with standard non-motion-based controller such as PS or Xbox

Console with motion-based controller such as Nintendo Wii or Xbox WITH a Kinect available

PC with standard non-motion controller

PC with motion controller

VR, please name which:

6. How frequently do you play games that utilise player movement?

A couple times per year

A couple times per month

About 1 time per week

2-4 times a week

5 or more times per week

7. Do you have any experience playing a musical instrument?

YES, state which:

NO – **Jump to question 10**

8. How many years have you played?

Less than 1

2-4 years

5-6 years

7+ years

9. How often do you practice?

A couple times per year

A couple times per month About once a week

2-4 times per week

5 or more times per week

10. Do you have any experience using music production software such as Ableton Live or Cubase?

YES

NO

11. How were you first introduced to musical concepts?

1 on 1 instrument lessons with a teacher/tutor

Group lessons in a class with a teacher

At home with parent/guardian

Self-Taught

Other, please name:

Never been introduced to musical concepts.

Section 1: Playing Experience

1. How intuitive was the interaction?

Not at all				Very
1	2	3	4	5

2. How responsive was the input on the sound in the game?

Not at all				Very
1	2	3	4	5

3. How challenging was the game?

Not at all				Very
1	2	3	4	5

4. Where did you place your attention?

Arranging sound

Reaching the end platform

I don’t know

5. How in control did you feel?

Not at all				Very
1	2	3	4	5

6. Did you reach the end platform?

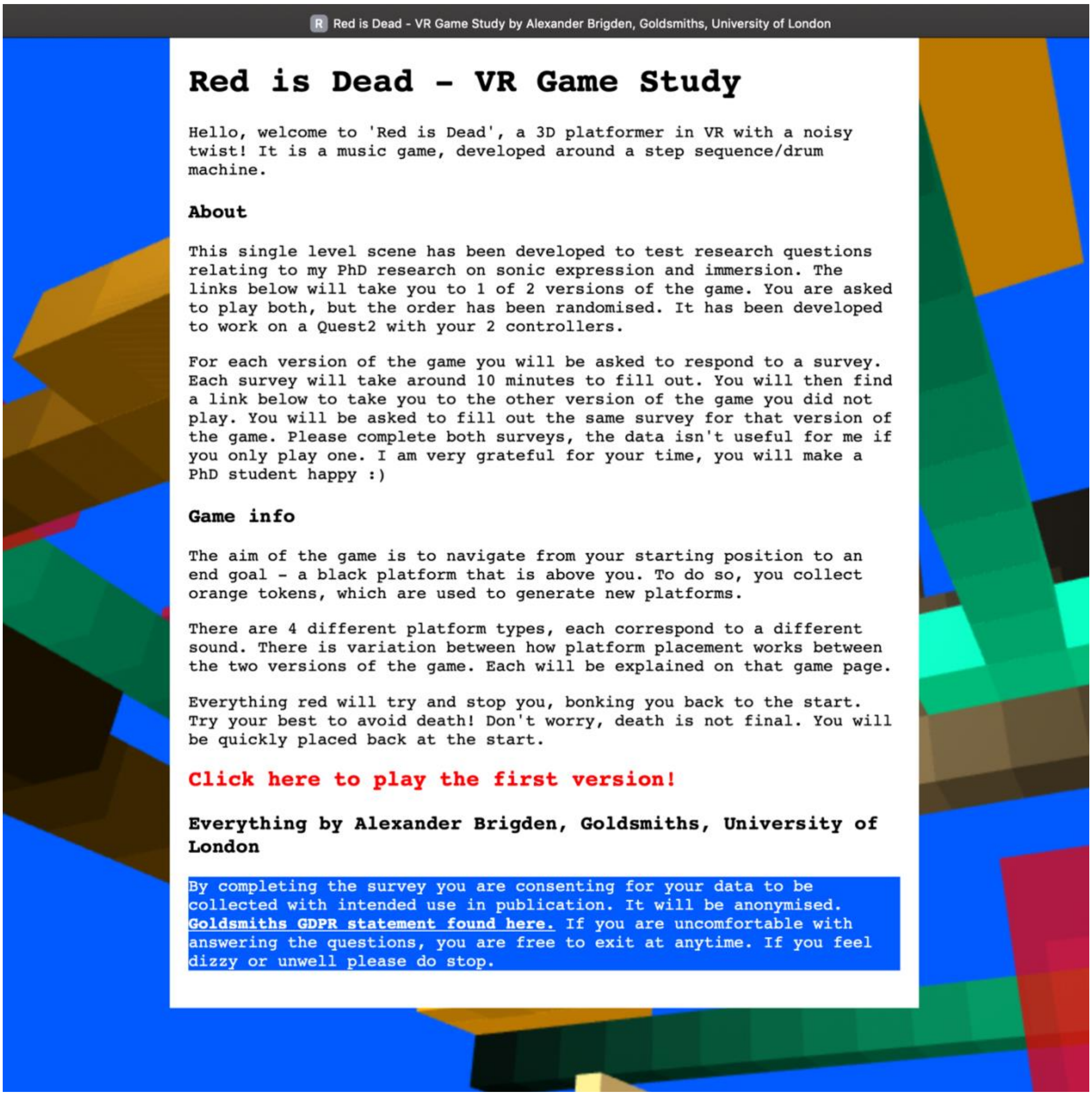
Yes

No

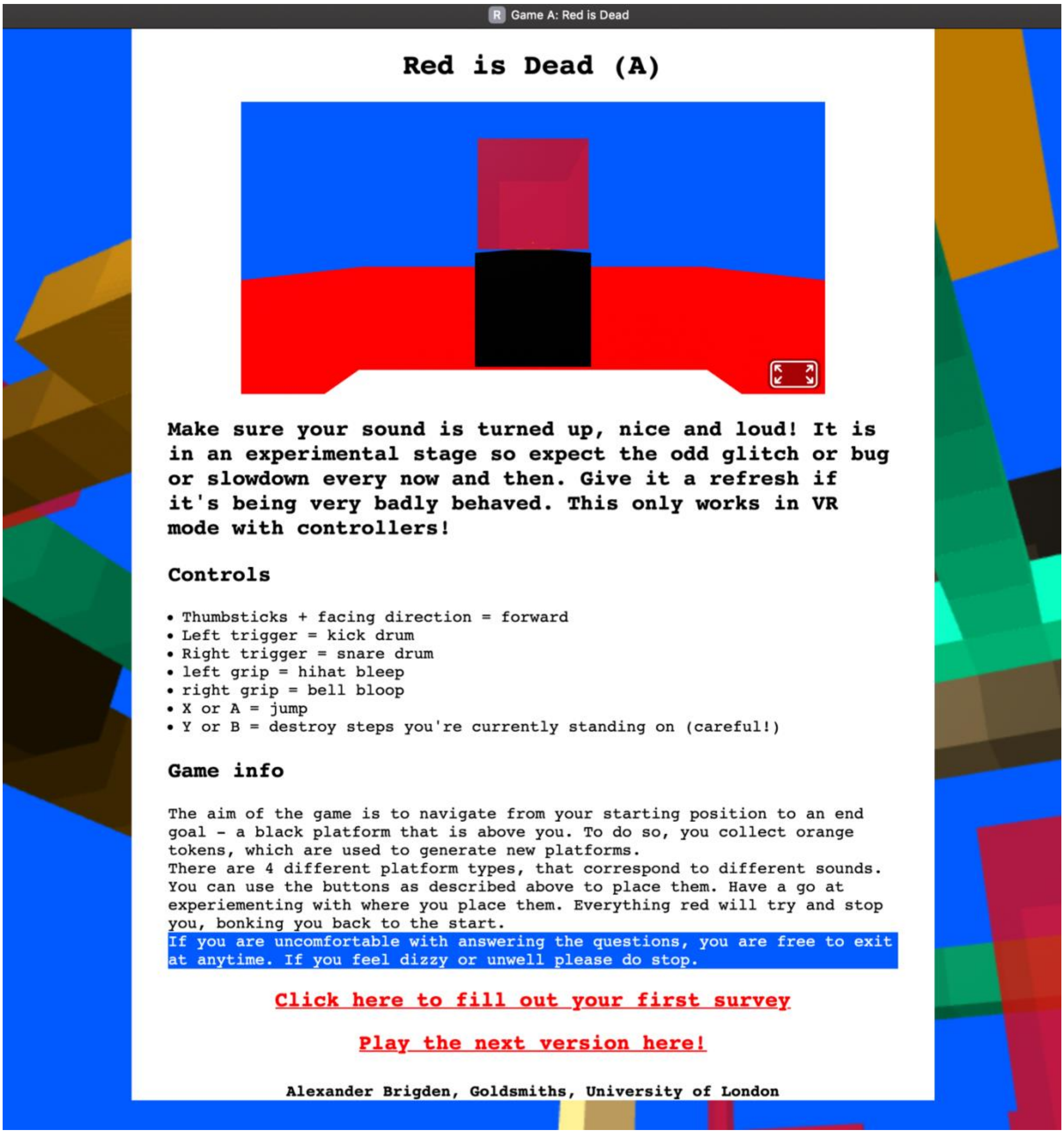
Section Two – IEQ / Immersion

Section Two was comprised of the 32 IEQ questions on a five-point Likert scale as per Jennet et al 2008 paper (Jennet el al., 2008).

Study Two Website Splash Page



Study Two Website Game A Page



Game A: Red is Dead

Red is Dead (A)

Make sure your sound is turned up, nice and loud! It is in an experimental stage so expect the odd glitch or bug or slowdown every now and then. Give it a refresh if it's being very badly behaved. This only works in VR mode with controllers!

Controls

- Thumbsticks + facing direction = forward
- Left trigger = kick drum
- Right trigger = snare drum
- left grip = hihat bleep
- right grip = bell bloop
- X or A = jump
- Y or B = destroy steps you're currently standing on (careful!)

Game info

The aim of the game is to navigate from your starting position to an end goal - a black platform that is above you. To do so, you collect orange tokens, which are used to generate new platforms.
There are 4 different platform types, that correspond to different sounds. You can use the buttons as described above to place them. Have a go at experiementing with where you place them. Everything red will try and stop you, bonking you back to the start.
If you are uncomfortable with answering the questions, you are free to exit at anytime. If you feel dizzy or unwell please do stop.

Click here to fill out your first survey

Play the next version here!

Alexander Brigden, Goldsmiths, University of London

Study Two Website Game B Page

Game B: Red is Dead

Red is Dead (B)

Make sure your sound is turned up, nice and loud! It is in an experimental stage so expect the odd glitch or bug or slowdown every now and then. Give it a refresh if it's being very badly behaved. This only works in VR mode with controllers!

Controls

- Thumbsticks + facing direction = forward
- X or A = jump

Game info

The aim of the game is to navigate from your starting position to an end goal – a black platform that is above you. To do so, you collect orange tokens, which are used to generate new platforms.
All you have to do is collect the tokens and avoid the red playheads. The token placement will happen when you pick them up.
If you are uncomfortable with answering the questions, you are free to exit at anytime. If you feel dizzy or unwell please do stop.

Please fill out your second survey here

Thank you for playing both of the games! :)

Alexander Brigden, Goldsmiths, University of London

Study Three Game Design Document

Rhythm Game Prototype

Working Title: Grail

Introduction

This 3D VR shooter uses musical arrangement as a gaming mechanic to generate 3D targets for the player to shoot and avoid. As music players corresponding target blocks appears. The players must shoot lasers at all targets that move towards them before they hit the player. The aim is for the player to reach the end of the music / level without taking too much damage from being hit by the targets. If the player does take too much damage, then the game prematurely ends.

Gameplay

The player is instructed to destroy all blocks before they hit the player. To do so they can fire different coloured lasers out of each of the Oculus Touch controllers. The controllers are visible in the game environment, and the laser's direction is visible with a semi-transparent sight, facing the perpendicular to the shoulder triggers on the controller. Each controller has a different coloured laser (Left is pink, right is yellow). The lasers can only destroy the corresponding-coloured targets. The targets are generated from a midi file of a music composition by this researcher. All the targets correspond to midi notes for the bassline. The pitch of the note determines the Y and Z position values, the duration of the note determines the z length of the target block. Once the player presses any button the game and music will start. As the music plays it moves through the midi arrangement sequentially. The midi notes alternate in different patterns in front of the player, from centre left and centre right. In the 'A' version of the game when the player shoots a target it plays that corresponding bass note on a bass synthesiser. The backing track for the A version contains a choral pad and drumbeat. In the 'B' version of the game it plays a sound effect when the block is hit by the player. The backing music contains the same arrangement featuring a choral pad, the bassline, and the drumbeat. If the player gets hit by a block, they take damage. Upon taking damage the players vision will start to flash red. Initially this will highly transparent. As the player takes damage the red becomes opaquer. The flashes are in metric time with the music. If the player takes too much damage the screen will remain red, and the music will stop. Additionally, if the player takes damage a sound effect will play each time they are hit. If the player reaches the end of the game (and the music) without taking too much damage they are awarded a grail!

Game Environment

The player's position is fixed with the game environment. The environment has a black background in all directions. In front of the player there are four evenly spaced grey lines that act as playheads (these represent a 'beat' relative to the musical arrangement). There is also a low-lit wall spaced a beat behind the last line. The left wall is pink, the right wall is yellow. The player can see coloured lasers facing away from their controllers (left is pink, right is yellow). When no button is pressed on the controller, they are faint in transparency. When the button is pressed, these turn solid. As the game starts music plays target blocks (that correspond with the bassline of the music) for the player to shot, appear out of the coloured walls. When a block appears it momentarily turns the whole wall solid in that colour. As a block approaches a playhead line on the floor, the line increases in opacity to white, before fading back to grey as it moves away. It does then for all four of the playheads. In the 'A' version of the game, the nearest playhead is the 'true' playhead that is synchronised to the backing music. When a note is shot and triggered over this position it is in time with the backing music and the original arrangement for the bassline. If it is shot a playhead early, it is a beat early compared to the original arrangement. In the 'B' version this is offset by a bar, meaning the 'true' playhead that is synchronised to the arrangement playback is the wall at the back of the level. When the players take damage, their vision will flash red in time with the music, becoming increasingly opaquer with more damage.

Target Generation

Targets are generated by a midi file of a single bassline (as part of a complete musical arrangement). The midi file was quantised and parsed into JavaScript via a Python script offline. The midi event position was converted into sequence, quantised by 16ths. Along with the midi note pitch value and a length value, quantised). Each event was object, with the entire sequence an array of these objects. This was export from python as txt which could be easily turned into a JavaScript variable and file.

Example sequence: var technoBassline = [{s: 0, p: 57, l: 4}, {s: 4, p: 59, l: 4},…];

Once there is an array of a musical sequence a simple ruleset could be devised containing conditional logic for where to place certain pitches, duration, changes of rotation/position based on bar position (A / B ternary form etc). This allowed for easy level arrangement as the musical arrangement becomes the level, the two are intrinsically connected.

The decision to do so offline, as opposed to reading direct from a midi file stored online, is to avoid processing the midi file and sequencing the blocks 'live' on the Oculus browser. Avoiding performance issues with timing/lag is imperative given the limited computational power we have on a Quest 2 (and given we have control limits regarding memory allocation and garbage collection in JavaScript/Browser).

Levels

There is a single level in the study prototype. As described above under the heading 'Target Generation' It is quite easy to generate varying levels. Given the minimal art style it would be easy to generate further levels. However, this is unnecessary for the study.

Winning

In order to win the game players must reach the end of the level. If they take too much damage and the music stops it is game over. The player will need to refresh the a-frame canvas to try again. There is no limit on how many attempts they can have.

Controls

Players can move their head around to move the camera orientation
Left Trigger = fire pink laser
Right Trigger = fire yellow laser

Interface

The game is presented in an A-frame canvas on a webpage. The webpage has the control guide and the game information on how to play. They are instructed to enter VR mode to start. The game does not work unless on an Oculus Quest device.

Art

The art design is deliberately clean and concise. It uses simple bold colours to differentiate between the different game environment elements, and the target blocks for the players to shoot. The corresponding laser are clearly coloured so players understand they can only shoot certain targets with certain lasers. It is deliberately minimal to not distract from the sound in the game that is being tested. What little choice regarding visual aesthetic, such as colour palette was made by this research to compliment the timbral qualities of the music presented.

Sound

The sound and musical elements to the game are all designed by this researcher. I made a deliberate choice of techno as a genre/style due to the (typically) relatively slow bpm, compared to other genres of dance music. The backing track has a bpm of 127. Stylistically the music strafes minimal and progressive techno, in homage to artists found on Border Communities (record label) such as Extrawelt, Nathan Fake, Petter, Ricard Tobar, and more. This style also lent itself to change the distribution of notes, progressively getting more complex as the music developed over the course of the full track. Additionally, this also allowed me to keep the instrumentation down to three synthesisers and a drum machine. For the 'A' version, the bass synthesiser is being triggered in browser, via MaxiInstruments. This was beneficial for triggering notes with accurate timing (given this is a rhythm game, timing accuracy is crucial to the playing experience). The remaining backing music exists as a sample being triggered once upon start. The sound effect for when the player takes damage is a short one-shot sample, originally synthesised by this research. In The 'B' Version the backing track contains the original bassline. The destruction of the target blocks is another one-shot sample, originally synthesised by this researcher.

Study Three Survey Example

1. Name:

(for referencing/matching questionnaires only, will be kept off record If a service like survey monkey is used will need to match surveys together)

Demographic Questions

2. Age:

18-29, 30-39, 40-49, 50-59, 60+

3. How Long have you played video games for?

Less than 1 week, Less than 1 month, Less than 1 year, Less than 5 years, 5 years +

4, **How frequently do you play?**

A couple times per year

A couple times per month A

bout one day a week

2-4 days per week

5 or more days per week

5. What devices do you usually play on? Tick multiple where applicable.

Phone/Tablet

Nintendo DS or other handheld game specific device

Console with standard non-motion-based controller such as PS or Xbox

Console with motion-based controller such as Nintendo Wii or Xbox WITH a Kinect available

PC with standard non-motion controller

PC with motion controller

VR, please name which:

6. How frequently do you play games that utilise player movement?

A couple times per year

A couple times per month

About 1 time per week

2-4 times a week

5 or more times per week

7. Do you have any experience playing a musical instrument?

YES, state which:

NO – **Jump to question 10**

8. How many years have you played?

Less than 1

2-4 years

5-6 years

7+ years

9. How often do you practice?

A couple times per year

A couple times per month About once a week

2-4 times per week

5 or more times per week

10. Do you have any experience using music production software such as Ableton Live or Cubase?

YES

NO

11. How were you first introduced to musical concepts?

1 on 1 instrument lessons with a teacher/tutor

Group lessons in a class with a teacher

At home with parent/guardian

Self-Taught

Other, please name:

Never been introduced to musical concepts.

Section 1: Playing Experience

1. How much did it feel like a musical experience?

Not at all				Very
1	2	3	4	5

2. How intuitive was the interaction?

Not at all				Very
1	2	3	4	5

3. How responsive was the input on the sound in the game?

Not at all				Very
1	2	3	4	5

4. How challenging was the game?

Not at all				Very
1	2	3	4	5

5. To what extent were you focusing on syncing your input to the backing music?

Not at all				Very
1	2	3	4	5

6. How synchronised did the audio-visual experience feel?

Not at all				Very
1	2	3	4	5

7. Which sensory feedback was the most useful in guiding your input?

Sound/by ear

Visual cues/by sight

8. To what extent did you use the flashing visual cues to guide your input?

Not at all				Very
1	2	3	4	5

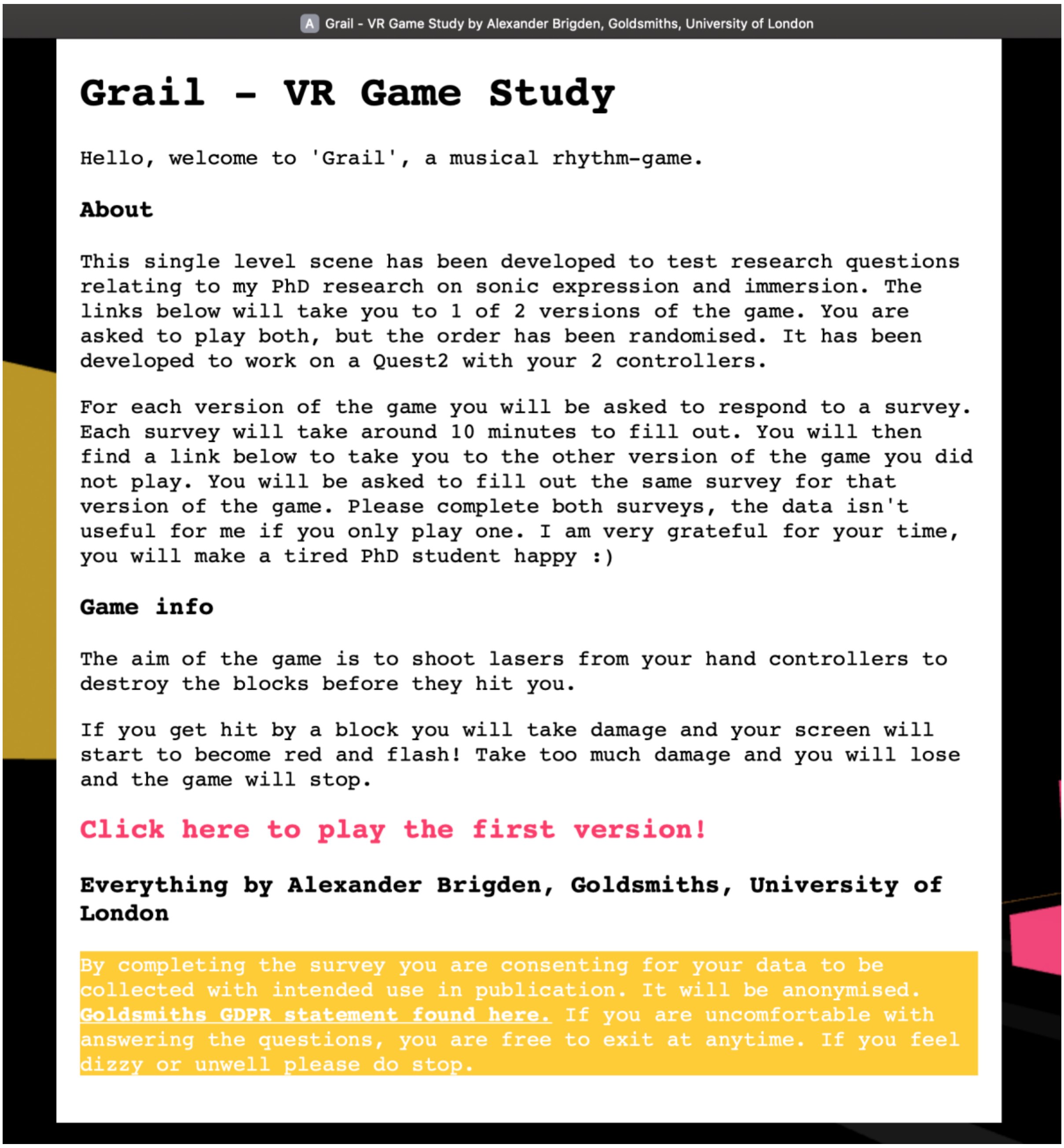
9. How much did missing a target/block and being hit break your focus?

Not at all				Very
1	2	3	4	5

Section Two – IEQ / Immersion

Section Two was comprised of the 31 IEQ questions on a five-point Likert scale (Jennet et al., 2008) and a 10-point Likert baseline.

Study Three Website Splash Page



Grail - VR Game Study by Alexander Brigden, Goldsmiths, University of London

Grail – VR Game Study

Hello, welcome to 'Grail', a musical rhythm-game.

About

This single level scene has been developed to test research questions relating to my PhD research on sonic expression and immersion. The links below will take you to 1 of 2 versions of the game. You are asked to play both, but the order has been randomised. It has been developed to work on a Quest2 with your 2 controllers.

For each version of the game you will be asked to respond to a survey. Each survey will take around 10 minutes to fill out. You will then find a link below to take you to the other version of the game you did not play. You will be asked to fill out the same survey for that version of the game. Please complete both surveys, the data isn't useful for me if you only play one. I am very grateful for your time, you will make a tired PhD student happy :)

Game info

The aim of the game is to shoot lasers from your hand controllers to destroy the blocks before they hit you.

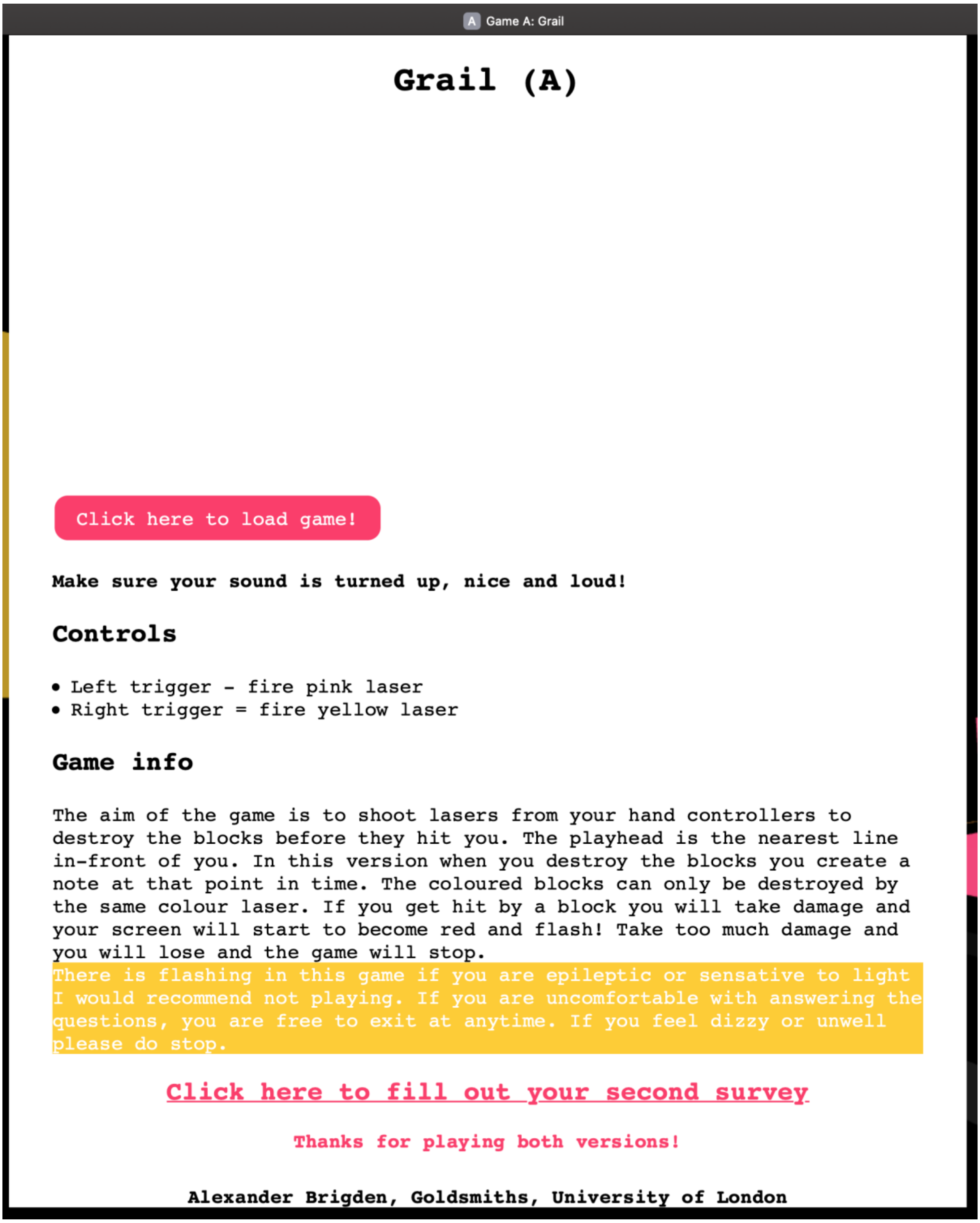
If you get hit by a block you will take damage and your screen will start to become red and flash! Take too much damage and you will lose and the game will stop.

Click here to play the first version!

Everything by Alexander Brigden, Goldsmiths, University of London

By completing the survey you are consenting for your data to be collected with intended use in publication. It will be anonymised. **Goldsmiths GDPR statement found here.** If you are uncomfortable with answering the questions, you are free to exit at anytime. If you feel dizzy or unwell please do stop.

Study Three Website Game A Page


Game A: Grail

Grail (A)

Click here to load game!

Make sure your sound is turned up, nice and loud!

Controls

- Left trigger - fire pink laser
- Right trigger = fire yellow laser

Game info

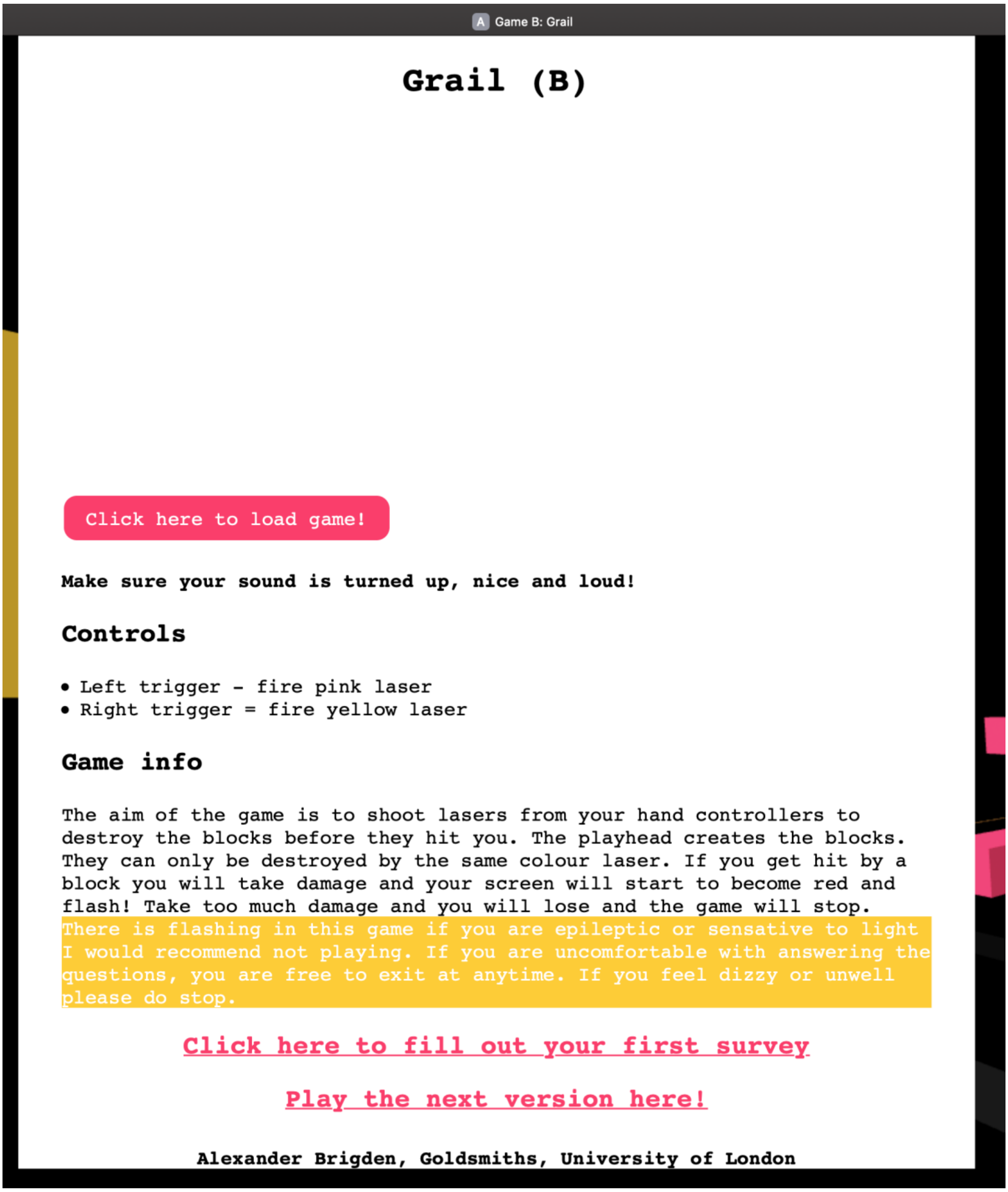
The aim of the game is to shoot lasers from your hand controllers to destroy the blocks before they hit you. The playhead is the nearest line in-front of you. In this version when you destroy the blocks you create a note at that point in time. The coloured blocks can only be destroyed by the same colour laser. If you get hit by a block you will take damage and your screen will start to become red and flash! Take too much damage and you will lose and the game will stop.
There is flashing in this game if you are epileptic or sensative to light I would recommend not playing. If you are uncomfortable with answering the questions, you are free to exit at anytime. If you feel dizzy or unwell please do stop.

Click here to fill out your second survey

Thanks for playing both versions!

Alexander Brigden, Goldsmiths, University of London

Study Three Website Game B Page


Game B: Grail

Grail (B)

Click here to load game!

Make sure your sound is turned up, nice and loud!

Controls

- Left trigger - fire pink laser
- Right trigger = fire yellow laser

Game info

The aim of the game is to shoot lasers from your hand controllers to destroy the blocks before they hit you. The playhead creates the blocks. They can only be destroyed by the same colour laser. If you get hit by a block you will take damage and your screen will start to become red and flash! Take too much damage and you will lose and the game will stop. There is flashing in this game if you are epileptic or sensative to light I would recommend not playing. If you are uncomfortable with answering the questions, you are free to exit at anytime. If you feel dizzy or unwell please do stop.

Click here to fill out your first survey

Play the next version here!

Alexander Brigden, Goldsmiths, University of London